

XX 26-SEP-2000; 2000EP-0308421.
 XX
 XX 29-SEP-1999; 99US-0156602.
 XX
 XX (PFIZ) PFIZER PROD INC.
 XX
 XX King KW, Madura RA, Rosey EL;
 XX
 XX WPI: 2001-309781/73.
 DR N-PSDB; AAS03285.
 XX
 XX New apoprotein antigens encoded by mhp3 gene from Mycoplasma
 PT hyopneumoniae useful as a vaccine for treating or preventing diseases
 PT caused by Mycoplasma hyopneumoniae -
 XX
 XX
 XX Claim 11; Page 17-18; 38pp; English.

XX The sequence is Mycoplasma hyopneumoniae MHP3 antigen. MHP3 antigen
 and its fragments are useful in manufacturing a vaccine for treating or
 CC preventing a disease or disorder in an animal, especially pig, caused
 CC by M. hyopneumoniae infection e.g. enzootic mycoplasma pneumonia.
 CC The mhp3-encoded proteins may be used as immunogens to generate
 CC antibodies which immunospecifically bind such an immunogen. The
 CC antibodies generated against the antigen are useful in diagnostic
 CC immunoassays, passive immunotherapy and generation of anti-idiotypic
 CC antibodies. Mhp3 proteins may also be used in immunoassays, e.g. to
 CC detect or measure in a biological sample from a vaccinated or
 CC potentially infected test animal the presence of antibodies to the
 CC antigen, and thus to monitor the immune response and/or to diagnose
 CC infection of the animal.
 XX
 XX Sequence 451 AA;

Query Match 99.0%; Score 2134; DB 22; Length 451;
 Best Local Similarity 99.8%; Pred. No. 8.6e-163;
 Matches 419; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 WDKETTEEKSADNQNOKITDVSKISGLVNERKSEIMAAKADANKHFLMAIYTAGTV 61
 DB 30 wdktctkeeksdnqkqtdvskisglvnerkseimaakadankhflmaiytagtv 89
 QY 62 NDNSNOSGWEAIQOLGALVGGETTSVDSSTAELEGKYSLSIANTNKNWVLSGFOHDAF 121
 DB 90 ndnsfngsweaiqqlgalvtgsetsvdsstaelegkyslsiantnknwvlsqfghdaf 149
 QY 122 TRWLKIPENKOLFTEKNIIIGIDMTDTEENVIPTRGYINLTYYKTEAGWLAGYANASFLA 181
 DB 150 trwlkipenkqlftekniililgldwtclenviptrgryinllykteegwlagyanasfla 209
 QY 182 KKEPSDPTKRSALVIGGISPAVTDPIAGYLAGIKAWNLKNSDKTKITTDKIEINLGF 241
 DB 210 kkepsdptkrsalviggispaavtdpiagylagikawnlksndkktktitdktieinlghd 269
 QY 242 VODSTKRELEQIASKDKPSTLLAVAGPLTEIFSDIIANONDRYLIGVDFQSLVYTKK 301
 DB 270 vgdctckrleqiaekdkpstllavagpllteifsdlianondryligvdfqslvytktk 329
 QY 302 NKFFTSILKNCIGYFVSLSLDYTKKSNRNAGFEFGKSKATYVLGIRKDRVDAIDNSL 361
 DB 330 nkfftsilknigysvlsldytkksnrnlagfefgksskatyvlglkdrvdaiadsel 389
 QY 362 EGNDRKLTAEISEAKKEFEERTKTIIPAEVYKTLIEIPMPKODKOESLDKITTIN 421
 DB 390 egndrkltaeiseakkefeektkcipaevyktlieipmpokgdkqesldkiltidin 449

RESULT 2
 AAU01860
 ID AAU01860 standard; Protein: 423 AA.
 XX
 AC AAU01860;

XX 07-SEP-2001 (first entry)
 DT
 XX
 XX Mycoplasma hyopneumoniae recombinant MHP3 antigen.
 DE
 XX
 XX MHP3; antigen; vaccine; enzootic mycoplasma pneumonia; mutant;
 KW antibody; immunoassay; immunotherapy; anti-idiotypic antibody; mutein.
 KW
 XX
 OS Mycoplasma hyopneumoniae.
 FH
 FH Key Location/Qualifiers
 FT Misc-difference 70
 FT Misc-difference /note= "Encoded by RGA"
 FT 417..422
 FT Misc-difference /note= "Encoded by AATACGATAT"
 FT
 XX
 XX EPI090995-A2.
 PN
 XX
 PD 11-APR-2001.
 XX
 XX 26-SEP-2000; 2000EP-0308421.
 XX
 XX 29-SEP-1999; 99US-0156602.
 XX
 XX (PFIZ) PFIZER PROD INC.
 XX
 XX King KW, Madura RA, Rosey EL;
 PI
 XX WPI: 2001-309781/73.
 DR N-PSDB; AAS03286.
 XX
 XX New apoprotein antigens encoded by mhp3 gene from Mycoplasma
 PT hyopneumoniae useful as a vaccine for treating or preventing diseases
 PT caused by Mycoplasma hyopneumoniae -
 XX
 XX
 XX Claim 4; Page 19-21; 38pp; English.

CC The sequence represents Mycoplasma hyopneumoniae a recombinant MHP3
 CC antigen lacking the first 28 amino acids (the putative signal sequence).
 CC MHP3 antigen and its fragments are useful in manufacturing a vaccine for
 CC treating or preventing a disease or disorder in an animal, especially
 CC pig, caused by M. hyopneumoniae infection e.g. enzootic mycoplasma
 CC pneumonia. The mhp3-encoded proteins may be used as immunogens to
 CC generate antibodies which immunospecifically bind such an immunogen. The
 CC antibodies generated against the antigen are useful in diagnostic
 CC immunoassays, passive immunotherapy and generation of anti-idiotypic
 CC antibodies. Mhp3 proteins may also be used in immunoassays, e.g. to
 CC detect or measure in a biological sample from a vaccinated or
 CC potentially infected test animal the presence of antibodies to the
 CC antigen, and thus to monitor the immune response and/or to diagnose
 CC infection of the animal.
 CC
 XX
 XX Sequence 423 AA;

Query Match 98.4%; Score 2120; DB 22; Length 423;
 Best Local Similarity 100.0%; Pred. No. 1e-161;
 Matches 416; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKDKTTEEKSADNQNOKITDVSKISGLVNERKSEIMAAKADANKHFLMAIYTAGTV 60
 DB 1 mdkctctkeeksdnqkqtdvskisglvnerkseimaakadankhflmaiytagt 60
 QY 61 VNDNSFNOSGWEAIQOLGALVGGETTSVDSSTAELEGKYSLSIANTNKNWVLSGFOHDA 120
 DB 61 vndnsfngsweaiqqlgalvtgsetsvdsstaelegkyslsiantnknwvlsqfghda 120
 QY 121 FTRWLKIPENKOLFTEKNIIIGIDMTDTEENVIPTRGYINLTYYKTEAGWLAGYANASFL 180
 DB 121 ftrwlkipenkqlftekniililgldwtclenviptrgryinllykteegwlagyanasfl 180
 QY 181 AKKEPSDPTKRSALVIGGISPAVTDPIAGYLAGIKAWNLKNSDKTKITTDKIEINLGF 240
 DB 181 akkepsdptkrsalviggispaavtdpiagylagikawnlksndkktktitdktieinlghd 240

Db 181 akfipdptkrtsaivvggispavtffiaaglyaglkawlnksdkkltctdkieinlgf 240
 QY 241 DVODSTKERLEQIASKDKPSTLLAVAGPLTEIFSDIIANQNDRYLIGVTDQSLVYTKT 300
 Db 241 dvodstkerleqiaaskdkpstllavaagplteifsdianqndrylignvtdqslvytkt 300
 QY 301 KNEFTSLIKNLGYSVFVSLDLYTKKSNRLAGFEFGKSAVYVLGIDRFVADIADTS 360
 Db 301 knkfftsllknlgysvfvsladlytkksnrlagfegfksatvylgldrfvadiads 360
 QY 361 LBNDDKKLTAEISAEKKEFEKTKTTPAEVAKTLEIPEMPKOPKQOESIDKL 416
 Db 361 legndkkltaleisaekekfeektktipaeevtktleipempdkpdkqesidkl 416
 RESULT 3
 AAR40856
 ID AAR40856 standard; Protein; 453 AA.
 AC AAR40856;
 XX
 DT 07-MAR-1994 (first entry)
 XX 43kd regression associated antigen.
 KM Regression associated antigen; tumour; immunotherapy;
 KW anti-idiotypic antibodies; antibodies; tumour regression.
 XX
 OS Mycoplasma hyorhinis.
 FH Key Location/Qualifiers
 FT Misc-difference 80 /note- "Tryptophan encoded by TGA, normal in
 FT Misc-difference 124 /note- "Mycoplasma hyorhinis."
 FT Misc-difference 165 /note- "Tryptophan encoded by TGA, normal in
 FT Misc-difference 165 /note- "Tryptophan encoded by TGA, normal in
 FT Misc-difference 344 /note- "Tryptophan encoded by TGA, normal in
 FT Misc-difference 344 /note- "Mycoplasma hyorhinis."
 XX
 PN US5242823-A.
 PD 07-SEP-1993.
 XX
 PE 07-MAR-1986; 86US-0837494.
 XX
 PR 07-MAR-1986; 86US-0837494.
 PR 16-SEP-1987; 87US-0097910.
 PR 11-DEC-1987; 87US-0131815.
 PR 04-JAN-1988; 88US-0138923.
 PR 16-MAR-1990; 90US-0474730.
 PR 02-OCT-1992; 92US-0956546.
 XX
 PA (ITGE-) INT GENETIC ENG INC.
 XX
 PI Fareed GC, Ghosh-dastidar P, Jar-how L, Sen A;
 XX
 DR WPI; 1993-295229/37.
 DR N-PDB; AAQ47816.
 XX
 PT DNA encoding a regression-associated antigen from M. hyorhinis -
 PT is used to obtain prods. for diagnosis, localisation and therapy
 PT of tumours
 XX
 PS Disclosure; Figure 3; 40pp; English.
 CC Regression associated antigens (RA's) are identified in material
 CC from neoplastic cells by their immunological reactivity with
 CC regression associated antibodies from the serum of patients

CC diagnosed as undergoing regression of a tumour. RA's can be used
 CC for tumour immunotherapy and for producing and purifying antibodies
 CC which can be used for tumour diagnosis, localisation and therapy.
 CC The antibodies can also be used for the production of
 CC anti-idiotypic antibodies which can also be used in immunotherapy.
 XX
 SQ Sequence 453 AA;
 Query Match 32.6%; Score 702; DB 14; Length 453;
 Best Local Similarity 38.8%; Pred. NO. 6.7e-48;
 Matches 170; Conservative 82; Mismatches 146; Indels 40; Gaps 14;
 QY 5 ETTKEKSAONOKQITDVSKISGLVNERKSEIIMAAKADANKHGLMAIVTAGTVQNDN 64
 Db 28 etdkegkll-----rlfd-----nsfvdrqaeleka-----knfdftvlltagvtqdk 73
 QY 65 SFNOSGWEAT---QOLGALTG-GEITSVDSTAELEKYSILANNKNVWVLSGFOHGD 119
 Db 74 sfngslweavlenydgletektndrvsqetnngselgkyknfngknwvlltgfgqg 133
 QY 120 AFTRWLKIPEN---KQLTEKNIILGIDW---TPDEVIPTRGRINTYKTEEAGWL 171
 Db 134 efpkflqtasngkkyssdliaekkvliavawdwlskedkdlkagnfisllykteeagfl 193
 QY 172 AGYANASFLAKKPPSPDKRSATVIGGDISPAVDFIAGYLAGIKAWNLSKDKTKITTT 231
 Db 194 ayyasakflaykfpndeakrtliapfgsgnagayrdfiaagflaglakynndpnackvlisd 253
 QY 232 DKIEINLGFVODSTKERLEQIASKDKPSTLLAVAGPLTEIFSDII--ANQNDRYLIGV 289
 Db 254 nlnldtqf--lsndktatfnglyvns--slvjpyvsgltsvvdalksnkdkyllgv 310
 QY 290 DPOOSLYYTKTKNKEFTSLIKNLGYSVFVSLDLYTKKSNRLAGFEFGK---KQATV 345
 Db 311 dtqgsklirpa-tvftfslethlgrtyevltdvllkedsrkfsgfslrlnpanav 369
 QY 346 YLGIDRFVADIATSLBGNDKLATTEAISAEKKEFEKTKTTPAEVAKTLEIPEM--PD 403
 Db 370 ykgsddfvgsnsvtadackvkqeflnaetadfkxqanpc-nyksvlgpmtlnd 428
 QY 404 KQPDKQOESIDKLITDIN 421
 Db 429 ndaknekainelkkin 446
 RESULT 4
 AAR67582
 ID AAR67582 standard; Protein; 432 AA.
 XX
 AC AAR67582;
 XX
 DT 06-OCT-1995 (first entry)
 DE Cancer metastasis protein.
 XX
 KM Cancer metastasis; clinical tests; detection.
 XX
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Misc-difference 55 /note- "corresponding codon TGA"
 FT Misc-difference 99 /note- "corresponding codon TGA"
 FT Misc-difference 140 /note- "corresponding codon TGA"
 FT Misc-difference 140 /note- "corresponding codon TGA"
 FT Misc-difference 319 /note- "corresponding codon TGA"
 FT Misc-difference 420 /note- "corresponding codon TGA"
 FT Misc-difference 424 /note- "corresponding codon TGA"
 FT Misc-difference 424 /note- "corresponding codon TGA"

```

XX JP06319559-A.
XX
XX 22-NOV-1994.
XX
XX 12-MAY-1993; 93JP-0144165.
XX
XX 12-MAY-1993; 93JP-0144165.
XX
XX (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.
XX
XX WPI: 1995-040317/06.
XX
XX N-PSDB; AAO79124.
XX
XX DNA encoding polypeptide involved in cancer metastasis - useful
XX PT for the study of the mechanism of metastasis and in clinical
XX PT tests
XX
XX Claim 1; Page 2; 13pp; Japanese.
XX
XX AAO79124 encodes AAR67582 a protein involved in cancer metastasis,
XX CC which may be used for the study of metastatic mechanisms, and for
XX CC clinical tests to determine the presence or absence of cancer
XX CC metastasis.
XX
XX SQ Sequence 432 AA;

Query Match 32.1%; Score 692.5; DB 16; Length 432;
Best Local Similarity 39.0%; Pred. No. 3.6e-47;
Matches 168; Conservative 81; Mismatches 141; Indels 41; Gaps 14;

OY 5 ETTKEKSDNKNKOITDVSKISGLVNERKSEIMAKADANKHFGIMNAVITAGTVNDN 64
DB 3 eckkegkll-----flfd-----nsfvdtrgaieleka-----knfdntvlltaagtvqdk 48
OY 65 SFNOSGWEAI-----QQLGALTG-GEITVSSTALEGKYSILANTNKNVWVLSGFQHD 119
DB 49 sfngsiweaveivlenhydqlektlnldrvsgeltnsgelldgkykfnlgnknvlltqfsggq 108
OY 120 AFTRWLKIPEIN-----KOLFTEKNIILIGIDW---TDENVIPTGRVYINLTKTEEAGWL 171
DB 109 efpkflkqtdsngkkysdliaekkvliavdwdlskedkdlkagfnlsilykteeagfl 168
OY 172 AGYANASFLAKKFPSPDPTKRSAIVIGGIGSPAVTDFIAGYLAGIRAMWLKNSDKRTKITT 231
DB 169 agyaasskflaykfnpdeakrtlapfggghagvtdflagflaglakyndnptgkvltssd 228
OY 222 DKIEINLGFVDVDTSTKERLEOIAISKDKPSTLLAVAGPLTEIFSDII--ANONDRYLIGV 289
DB 229 mnlndtgrf-lsndktatfinglvnks--slvlpvagsltsavdalikksnkdtkyllgv 285
OY 290 DTDOSLVYTKTKNKPFETSIILKNLGYSVFSVLSDLVTKKSNRNLAGFERGK---KSATV 345
DB 286 dtdgskflfpa-tvfftslekhgrtyvldtvlwkksedskflsgfsfkltpnanatv 344
OY 346 YLGIKDRFVDIADTSLSEGNCKKLATLTAISEAKKEFEKTKTTPAEVVKTLTEIPEM---P 402
DB 345 ykglisd-dflgvsnstvadaadvkageflneatadtkkgqanpt-nyksvlgipmlind 403
OY 403 DKOPPKQOESL 413
DB 404 ndakdnkasl 414

RESULT 5
AAP93343
ID AAP93343 standard; protein; 448 AA.
XX
XX AAP93343;
AC
XX
XX 01-AUG-1990 (first entry)
XX

```

```

DE Gene encoding the 43 kd regression-associated antigen (RA Ag) of
DE Mycoplasma hyorhinis.
XX
XX Tumour regression-associated antigens (RA Ag); in vivo imaging;
XX KW therapy monitoring; cancer therapy; Mycoplasma hyorhinis.
XX
XX Mycoplasma hyorhinis.
XX
XX Key Location/Qualifiers
XX FT Peptide 1..25
XX FT Protein 26..448
XX
XX EP308265-A.
XX
XX 22-MAR-1989.
XX
XX 16-SEP-1988; 88EP-0308625.
XX
XX 11-DEC-1987; 87US-0131815, US-097910.
XX
XX (ITGE-) INT GENETIC ENG INC.
XX
XX Fared GC, Sen A, Ghosh-Dastidar P, Liu A, Lee JH.
XX
XX WPI: 1989-087638/12.
XX
XX N-PSDB; AAN90684.
XX
XX Tumour regression-associated antigens and antibodies -
XX PT used in diagnostic tests, monitoring course of therapy and for
XX PT therapy in cancer patients
XX
XX PS Disclosure; : 56pp; English.
XX
XX Regression-associated antigens may be used in diagnostic tests, eg in
XX CC vivo imaging, for monitoring the course of therapy or for therapeutic
XX CC purposes, eg active immunisation protocols in cancer patients or drug
XX CC delivery systems by binding the drug to monoclonal or monospecific
XX CC polyclonal Ab showing specific immunoreactivity with the Ag.
XX
XX SQ Sequence 448 AA;

Query Match 29.8%; Score 642.5; DB 10; Length 448;
Best Local Similarity 38.0%; Pred. No. 3.9e-43;
Matches 166; Conservative 83; Mismatches 145; Indels 43; Gaps 18;

OY 5 ETTKEKSDNKNKOITDVSKISGLVNERKSEIMAKADANKHFGIMNAVITAGTVNDN 64
DB 28 eckkegkll-----flfd-----nsfvdtrgaieleka-----knfdntvlltaagtvqdk 73
OY 65 SFNOSGWEAI-----QQLGALTG-GEITVSSTALEGKYSILANTNKNVWVLSGFQHD 119
DB 74 sfngsiweaveivlenhydqlektlnldrvsgeltnsgelldgkykfnlgnknvlltqfsggq 133
OY 120 AFTRWLKIPEIN-----KOLFTEKNIILIGIDW---TDENVIPTGRVYINLTKTEEAGWL 171
DB 124 efpkflkqtdsngkkysdliaekkvliavdwdlskedkdlkagfnlsilykteeagfl 193
OY 172 AGYANASFLAKKFPSPDPTKRSAIVIGGIGSPAVTDFIAGYLAGIRAMWLKNSDKRTKITT 231
DB 194 agyaasskflaykfnpdeakrtlapfggghagvtdflagflaglakyndnptgkvltssd 253
OY 222 DKIEINLGFVDVDTSTKERLEOIAISKDKPSTLLAVAGPLTEIFSDII--ANONDRYLIGV 289
DB 254 mnlndtgrf-lsndktatfinglvnks--slvlpvagsltsavdalikksnkdtkyllgv 310
OY 290 DTDOSLVYTKTKNKPFETSIILKNLGYSVFSVLSDLVTKKSNRNLAGFERGK---KSATV 345
DB 311 dtdgskflfpa-tvfftslekhgrtyvldtvlwkksedskflsgfsfkltpnanatv 367
OY 346 YLGIKDRFVDIADTSLSEGNCKKLATLTAISEAKKEFEKTKTTPAEVVKTLTEIPEM-PDK 404
DB 368 ykglisd-flgvsnstvadaadvkagefln-etadtkkgqanpt-nyksvlgipmlind 424

```


OY 405 QPDKQGESLKLITDIN 421
 Db 425 dakhmekalnelikkln 441

RESULT 6

AAV05332
 ID AAV05332 standard; peptide; 428 AA.

AC AAV05332;

DT 25-JUN-1999 (first entry)

DE Inflammatory cytokine inducer.

KW Inflammatory cytokine inducer; thrombocytopenia; therapy.

OS Synthetic.

PN JP11089582-A.

XX 06-APR-1999.

PF 19-SEP-1997; 97JP-0273649.

PR 19-SEP-1997; 97JP-0273649.

XX (CHUS) CHUGAI PHARM CO LTD.

XX WPI: 1999-281057/24.

DR N-PSDB; AAX3847.

XX New inflammatory cytokine inducer gene and polypeptide - useful for

PT treatment of thrombocytopenia

XX Claim 3; Page 16-17; 22pp; Japanese.

CC This sequence represents the inflammatory cytokine of the

CC invention. The inflammatory cytokine can be used in a drug, which is

CC useful for the treatment of thrombocytopenia.

XX Sequence 428 AA;

Query Match 17.6%; Score 379.5; DB 20; Length 428;

Best Local Similarity 26.7%; Pred. No. 4.6e-22; Mismatches 156; Indels 71; Gaps 17;

Matches 115; Conservative 88; Mismatches 156; Indels 71; Gaps 17;

31 NERKSEIMAAKADANKH-----FGLNMAIVTAGTVNDNSFNQSGWEAI 74

27 ndesnlstfkekdskyttnangkvvnaelkikpylltdegklddsfngsafel 86

OY 75 QOLGALTGEELISVDSSTLEEGKYSLSANTKNVWVLSGFHGDFTKMLKIPENKOLF 134

Db 87 klnkqctglelnnveps-snfesaynsalsagkhivngfkhqgsikgy--ldahneel 143

OY 135 TEKNIIILGIDMTDFENVIPTGRYINLUTYKTEEGWLAGANSPPLAKKPPSDPTRSAI 194

Db 144 etnqkllidldt-dleeykw--fyslqfnlkkesafctgylaswseq--deskriva 197

OY 195 VTGGISPAVTFEIAAGYLAGIKAMNLSNDKKTI-TTDKIEINLGFVDSTKERLEQ 253

Db 198 sfsggsfpvvtfnegfakglllyngkh--ksklyhtspvkidsqf-----tagekmt 250

OY 254 IAS-----KDPSTLLAVAGPLTEIFSDIIANONDRYLIGVTDQSLVYTTTKNK 303

Db 251 vlnnvlsstpadkynphvllsvaspat--felvrlankqylyvosdgmml--qdkdr 306

OY 304 FPTSILKNIGSYVSVLSLDL-----YTKKSNSRNLAGFEFGKKSATVYLGTRDRF 353

Db 307 lltsvtkhkhqavetlidlillekeegykyprvvkdkkadkshfqtg-----kek 358

OY 354 VDIADTSLGENDKKL-ATEAISSEAKKEFEKTKTIPAEVYRKTLEIPMPD-KQPKQOE 411

Db 359 igvaenhfsnteegaklnkikemfke---lpedfvykynsdkaikdgkldnvs 414

OY 412 SLDKLITDIN 421

Db 415 rleallsaln 424

RESULT 7

AAW22727
 ID AAW22727 standard; Protein; 429 AA.

AC AAW22727;

DT 26-FEB-1998 (first entry)

DE Membrane protein M161Ag.

KW M161Ag; membrane protein; leukaemia; treatment.

OS Homo sapiens.

PN Key Location/Qualifiers

FT Modified-site 121 /note- "selenocysteine"

FT Modified-site 163 /note- "selenocysteine"

FT Modified-site 185 /note- "selenocysteine"

FT Modified-site 348 /note- "selenocysteine"

FT Modified-site 358 /note- "selenocysteine"

XX JP09157295-A.

PD 17-JUN-1997.

XX 05-DEC-1995; 95JP-0344504.

PR 05-DEC-1995; 95JP-0344504.

XX (KAGA-) KAGAKU GIYUTSU SHINKO JIGYODAN.

XX WPI: 1997-369470/34.

DR N-PSDB; AAT75133.

XX Membrane protein M161Ag - useful in development of leukaemia

PS treating agent

XX Claim 1; Fig 1; 6pp; Japanese.

CC This sequence is a membrane protein designated M161Ag. Oligonucleotide

CC probes for the cDNA were designed from putative N-terminal M161Ag peptide

CC sequences isolated from P39 cells. M161Ag can be used in the development

CC of a leukaemia treating agent.

XX Sequence 429 AA;

Query Match 17.3%; Score 372; DB 18; Length 429;

Best Local Similarity 26.8%; Pred. No. 1.9e-21; Mismatches 167; Indels 64; Gaps 18;

Matches 117; Conservative 88; Mismatches 167; Indels 64; Gaps 18;

OY 13 ADNOKOIT---DVSKISGLVNERKSEIMAAKADANKHFGLNMAIVTAGTVNDNSFNQ 68

Db 26 andesnlstfkekdskyttnangkvvnael-----lklypvlltdegklddsfng 80

OY 69 SGWEAIOQLGALTGETTSVDSSTLEEGKYSLSANTKNVWVLSGFHGDFTKMLKIP 128

Db 81 satealalnkhqvtglelnnveps-snfesaynsalsagkhixvlnghqgsikgy--ld 137

OY 129 ENKOLFTEKNIIILGIDWTDTEENVIPTRGYINLFTKTEBAGWLAVANASFLAKKPPSDP 188
 Db 138 ahrreelernqlkligldf-dleteyxx--fyslqfinkesaltgylalasxseq----de 191
 OY 189 TKRSALVIGGGSPAVTDEIAGYIKAMNLSKNSDKTKKI-TTDKIEINLGEFDQDST 247
 Db 192 srtvvaasfvggafpyvtctfnegfakyplyngkh--ksaklyhtspvryldsgf-----ta 244
 OY 248 KERLEQIAS-----KDRPSTLLAVAGPLTEIFSDIINQNDRLVIGDTPQSLV 297
 Db 245 gekmvtlnnvlnstpadkkyphvllavagpat--fevtlpmkgyvlygvsdsgm- 301
 OY 298 TTKTKKFTSILKNGYSFVSLSL-----YTKKSRLNLAGFEFGKKSATVYL 347
 Db 302 -gdkdrilltsvllkhlkgavycllqlillekeegykyrvvkkkkaakxshfgt----- 354
 OY 348 GKDRFVADIADTSLGNDKKL-ATEAISPAKKEFEKTKTIPAEVRKLTLEPEMPD-KQ 405
 OY 355 --kexxiyvaaenqfnteeqaklnmkleaklmfke----lpedivkylnsdkaikdgnk 408
 OY 406 PDKQESLIDKLITDIN 421
 Db 409 ldnveerleallisaaln 424

RESULT 8

AAV81632
 ID AAV81632 standard; Protein: 351 AA.

AAV81632;

24-MAY-2000 (first entry)

Streptococcus pneumoniae type 4 protein sequence #132.

KW Streptococcus pneumoniae; vaccine; screening; protein antigen;
 KW antibacterial; antinflammatory; meningitis; infection; diagnosis;
 KW pneumococcal disease.

XX Streptococcus pneumoniae.

XX WO200006737-A2.

XX 10-FEB-2000.

XX 27-JUL-1999; 99WO-GB02451.

XX 27-JUL-1998; 98GB-0016337.

XX 19-MAR-1999; 99US-0125164.

XX (MICR-) MICROBIAL TECHNICS LTD.

XX Gilbert CFG, Hansbro PM;

XX WPI: 2000-195300/17.

XX New Streptococcal protein, useful as a vaccine, for diagnosis of
 PT pneumococcal diseases and for screening agents capable of antagonizing
 PT or inhibiting expression of the protein

XX Claim 1; Page 85; 108pp; English.

XX AAV81501 to AAV81679 represent specifically claimed protein sequences
 CC isolated from Streptococcus pneumoniae. AAA05407 to AAA05590 represent
 CC specifically claimed nucleotide sequences isolated from S. pneumoniae.
 CC The sequences have antibacterial and antinflammatory properties.
 CC The protein sequences, and fragments of them, are useful as immunogens
 CC and/or antigens. The nucleotide sequences can be used in vaccines and in
 CC diagnostic assays. The proteins and nucleotides can be useful for the
 CC detection and diagnosis of S. pneumoniae. The protein sequences are also
 CC useful for screening an agent capable of antagonizing, inhibiting or
 CC interfering with the function or expression of the proteins in which the
 CC agent is useful for treatment or prophylaxis of S. pneumoniae infection

CC and meningitis. AAA05591 to AAA05614 represent primers used in the
 CC exemplification of the present invention.
 XX
 XX Sequence 351 AA;

Query Match 9.6%; Score 206; DB 21; Length 351;
 Best Local Similarity 25.5%; Pred. No. 2.9e-08;
 Matches 98; Conservative 57; Mismatches 146; Indels 84; Gaps 19;

OY 33 RKSEIMAAKADANKHFGIMNAIVTAGTVNDSEFNSQSMEMAIQQLGA---LTGGEITSDV 89
 Db 25 rrsrnaasssdvk---tkaalvtctgyvddksfngsaweglgavgkehnlslskdngfityf 80
 OY 90 SSTAELE-----GKYSLSLANTNNKWWVLSCFQHDADFTRLKIPENKOLFTEKNII 140
 Db 81 qtsaeadyannlqgaagynllfgy-----gfalmsa-----vdaakehdlnyv 126
 OY 141 ILGIDWTDTEENVIPTRGYINLFTKTEBAGWLAVANASFLAKKPPSDPTKRSALVIGGI 200
 Db 127 llddvlykdqkuv-----asvlfadneeagylagvaak-----tlckkqvgfvygl 171
 OY 201 -SPAVTDEIAGYIKAMNLSKNSDKTKITTDKIEINLGEFDVODTSTKERLEQIASDK 259
 Db 172 esevisrteagfkagv-----asvdpslkvydvag-sfg-daakgkclaaagyaagad- 223
 OY 260 PSTLLAVAGPL-TEIFSDIINQNDR-----YLIGVDTPQSL--VYT---KTKKPF 306
 Db 224 -lyvqvaggygagvfaaakslnesrpenekyvwlygvdrdgaagkykskdgkeanfylv 281
 OY 307 SILKNGYSFVSLSLDLYTKKSNSRLNLAGFEFGKKSATVYLGIKDRFVADIADTSLGNDK 366
 Db 282 slkkgvgtlvxdl-----snkaergefpggq--Vlyvslkdkgydlatvnlseegk 330
 OY 367 KLTEAISEAKKEFEKTKTIPAE 391
 Db 331 k-----avedakaklidsavkvpeks 351

RESULT 9

AAW55066
 ID AAW55066 standard; Protein: 328 AA.

AAW55066;

02-OCT-1998 (first entry)

Streptococcus pneumoniae SP007 protein.

KW Streptococcus pneumoniae; antigen; vaccine; infection; diagnosis;
 KW detection; pneumonia; otitis media; meningitis.

XX Streptococcus pneumoniae.

XX WO9818930-A2.

XX 07-MAY-1998.

XX 30-OCT-1997; 97WO-US19422.

XX 31-OCT-1996; 96US-0029960.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Chol GH, Hromocky A, Johnson LS, Kunsch CA;

XX WPI: 1998-272224/24.

XX N-PSDB: AAV27326.

XX Nucleic acid encoding antigenic peptide(s) from Streptococcus
 PT pneumoniae - or their epitope-containing fragments, useful in
 PT protective or therapeutic vaccines, and for diagnosis

Search completed: July 18, 2002, 15:32:20
Job time: 187 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2002, 15:39:55 ; Search time 33.42 Seconds
(without alignments)
1253.726 Million cell updates/sec

Title: US-09-676-249A-4
Perfect score: 2135
Sequence: 1 MMDETTKEKSKADNONKQI.....KQPDKQKSLDKLTIDNNL 423

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Minimum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	552	25.6	461	2	B90555	ABC transporter xy
2	203.5	9.4	350	2	F86804	basic membrane pro
3	202	9.4	350	2	G95097	lipoprotein [impor
4	198	9.2	374	2	C97865	conserved hypotet
5	192	8.9	357	2	D96986	probable lipoprote
6	181	8.4	353	1	H71340	membrane lipoprote
7	180.5	8.4	350	1	C70009	ABC transporter (l
8	178.5	8.3	359	2	F72418	basic membrane pro
9	177	8.2	341	2	E70147	basic membrane lip
10	173	8.0	357	2	AH1610	CD4+ T cell-stimul
11	168.5	7.8	516	2	C82946	hypothetical prote
12	168	7.8	357	2	AD1248	CD4+ T cell-stimul
13	165.5	7.7	360	2	H70147	basic membrane pro
14	162.5	7.5	525	2	C82914	conserved hypotet
15	161	7.5	339	2	F70147	basic membrane pro
16	152.5	7.1	591	2	D64204	membrane lipoprote
17	150	7.0	524	2	D82944	hypothetical membr
18	146	6.8	353	2	G70147	basic membrane pro
19	140	6.5	379	2	H75318	membrane lipoprote
20	139.5	6.5	337	2	B97374	Delonococcus radiod
21	139.5	6.5	337	2	AH2591	membrane lipoprote
22	136	6.3	547	2	E29504	mercuryl (IT) reduct
23	134	6.2	539	2	D82886	conserved hypotet
24	132	6.1	326	2	G95857	hypothetical prote
25	131.5	6.1	1558	2	B71603	RESA-H3 antigen PR
26	128	5.9	657	2	S73428	probable lipoprote
27	126	5.8	349	2	F84246	hypothetical prote
28	125.5	5.8	763	2	A82863	hypothetical prote
29	123.5	5.8	1223	2	E88451	protein K10D2.1 [1

30	123.5	5.7	626	2	C25035	colicin Ia - Esche
31	122.5	5.7	322	2	F84236	ABC transporter (l
32	122.5	5.7	2346	2	T13829	Tpr homolog - frui
33	120	5.6	1546	2	G90603	lipoprotein [impor
34	119.5	5.5	384	2	I40867	hypothetical prote
35	119.5	5.5	2285	2	T12796	probable transglyc
36	118.5	5.5	556	2	H82301	peptide ABC transp
37	118.5	5.5	1140	2	S73786	hypothetical prote
38	118.5	5.5	1963	2	B98002	IGA-specific metal
39	118	5.5	1031	2	C81302	probable type I sl
40	117	5.4	798	2	S62791	probable lipoprote
41	117	5.4	1959	2	AG1085	hypothetical prote
42	116.5	5.4	396	2	H97500	hypothetical prote
43	116.5	5.4	396	2	AD2719	conserved hypotet
44	115.5	5.4	350	2	F70139	exported protein (
45	115	5.3	282	2	C97271	probable xylanase/

ALIGNMENTS

RESULT 1	
B90555	ABC transporter xylase-binding lipoprotein [imported] - Mycoplasma pulmonis (strain U
C:Species: Mycoplasma pulmonis	
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001	
C:Accession: B90555	
R:Chamand, I.; Hellig, R.; Ferris, S.; Barde, V.; Samson, D.; Galisson, F.; Moszer,	
Nucleic Acids Res. 29, 2145-2153, 2001	
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma p	
A:Reference number: A99512; M0ID:21267165; PMID:11353084	
A:Accession: B90555	
A:Status: preliminary	
A:Molecule type: DNA	
A:Residues: 1461 <KUR>	
A:Cross-references: GB:AL445566; PID:g14089760; PIDN:CAC13519.1; GSPDB:GN00153	
A:Experimental source: strain UAB CTIP	
C:Genetics:	
A:Gene: MYPU_3460	
A:Genetic code: SGC3	
Query Match	25.6%; Score 552; Db 2; Length 461;
Best Local Similarity	31.2%; Pred. No. 5; Ie-27;
Matches 133; Conservative 82; Mismatches 161; Indels 50; Gaps 11;	
OY 13 ADNONKQIT--DYSKISGLYNER-----KSEIMAAKADANKHGLMMAITAGCTV 61	
DB AONPNKTNNSLDSKRIIDLLOSKEVETETOKIVENKIKQASLETOK---VLLTTADGNI 82	
OY 62 NDNSFNQSGWGAIOQLALNGEITS-----VDSSTALEEGKYSLSANTNK 107	
DB DKSFNQVYESQKTLDPFDKATKSONKEAENOKHDKNDYNSVVKLEQNYKVALDRGY 142	
OY 108 NWVVLISGFQHGDAFTRMKIPENKQLETKNIILIGIDWTDEN-VIPGRYINLYKTE 166	
DB TTWILTFQOQNEIENFLNDBNNLRPEKKNVKIIGVDAMPNANSKIPQSLSLLEKTE 202	
OY 167 EAGGLAIGYANASFLAKFPSPDPTKRSALVIGGGISPAVTPFIAGYLAGICIAWN--LKNSD 224	
DB EAGQAGVYASADLGTGYANNAEKRAISAFGGGDFAGVTFELNFFEGIGIANWSEANEN 262	
OY 225 KRTKTTDKIEINFGFVODSTFKERIEQJASRKPSLTLAVAGPLREIFSDIT---ANQ 281	
DB KRYATVSENLVDLGF-IPNAEKNEVSNVVECTKSITSLPVACPFVGVVVDYLRKTFSD 321	
OY 282 NDRYLISVDTQSLVYTKRKKEFTSLKNGYSFVSLSDLTY-----KKSNSRNL 333	
DB EDRIYVGVDTQSLFNTDSKRFTSLVKNIAFPVYQIILALLTKDESVYLRKGNOKFL 381	
OY 334 AGFPGKKSATVYIGRIDRVADIDTSLGENDKLALEAISEA--KKEFEKTTITPAEE 391	
DB 382 GS--NPKNLVLRGLAKFAFVNITKSHVKESIKTQADTSIOKALDKWANNPNNSKRIKEM 438	

OY 392 VKTLE 397
DB 439 TNGDLE 444

RESULT 2
F86804

basic membrane protein A [Imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: F86804
R:Botolin, A.; Winkler, P.; Manger, S.; Jallion, O.; Malarne, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s8
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: F86804

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-350 <STO>
A:Cross-references: GB:AE005176; PID:g12724428; PIDN:AAK05536.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: bmpa

Query Match 9.4%; Score 203.5; DB 2; Length 350;
Best Local Similarity 24.7%; Pred. No. 1.9e-05;
Matches 97; Conservative 63; Mismatches 157; Indels 75; Gaps 17;

OY 20 ITDVKISGLVNERKSEIMAAKADANKHFGMLMAIVTAGTVNDNSFNQSGWEALQOLGA 79
DB 13 LASVAVLAGCRSHDAGSGKATD-----LKAALVTEIGVNDSSFNQSAWEGLOSMCK 66
OY 80 LTG-----GEITSVDSSTAELGKYSLLANTKNVVLGSGFHGDATFRLKIPENKOLF 134
DB 67 ENMLKKGCTYTFQNSASDYYTNNSAQCGYKLLFGFGLQDPTSAAR---NN--- 120
OY 135 TEKNTIIGIDMTDENYIPGRYINLTYTEAGWLAGYANASFLAKKFPSPDKRSAT 194
DB 121 PKNFVIYDSVTKDQKNV-----ASATFADNESAYLAGVAAR-----ATKTKNI 165
OY 195 -VIGGIGSPAVTDFAGTAGICAKNNLNKSDKTKITTDKIEINLGFVODPTSKERLEQ 253
DB 166 GFTGMDQSDVITFEKGYEAGAKSVN-----PDIKVDVQYAGSFSDAAKKTTAA 215
OY 254 IASKDKPSTLLAVAGPL-TEIFSDIIA-----NONDR-YLIGVDTQSLV-YT---KT 300
DB 216 AMYGAGDDVVYGCAGVGTVSEAKALNSTKNEADKWVIVGDDQDEILCKYKSKDKD 275
OY 301 KKKFTSLIKNLGYSVSVLSDLYTKKSNRNLAG--FEFGKKSATVYLGIKDRFVDIAD 358
DB 276 SNFVLVSTIKVGV---NVVKDIADKTRDKGKPGGTIVYDLKNGGVNLGL-----DSAN 326
OY 359 TSLGNDKKLATEATSEAKKEFEKTKTTPAE 390
DB 327 SEIK-----DAVAKAKADIIDGKITVPSK 350

RESULT 3
G95097

lipoprotein (Imported) - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C:Accession: G95097
R:Tellet, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Held
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfle,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: G95097

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-350 <KUR>
A:Cross-references: GB:AE005672; PIDN:AAK74976.1; PID:g14972319; GSPDB:GN00164; TIGR:
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: Sp0845

Query Match 9.4%; Score 202; DB 2; Length 350;
Best Local Similarity 25.7%; Pred. No. 2.4e-05;
Matches 98; Conservative 55; Mismatches 145; Indels 84; Gaps 19;

OY 33 RKSEIMAAKADANKHFGMLMAIVTAGTVNDNSFNQSGWEALQOLGA---LTGGEITSVD 89
DB 25 RSSRNAASSDVK-----TKAALVDTGVDKSFQNSAMEGLQAMGKEHNLKDNKGFYF 80
OY 90 SSTALE-----GKYSLLANTKNVVLGSGFHGDATFRLKIPENKOLFTEKNI 140
DB 81 OSTSEADYANMLQOAGSNLIFGV-----GFLNNA-----VKDAKKEHLDLNVY 126
OY 141 ILGIDMTDENYIPGRYINLTYTEAGWLAGYANASFLAKKFPSPDKRSATYIGGI 200
DB 127 LIDDIYKQKNV-----ASVTFADNESGYLAGVAAR-----TTKQYGVGCI 171
OY 201 -SPAVTDFIAGTAGICAKNNLNKSDKTKITTDKIEINLGFVODPTSKERLEQIASRK 259
DB 172 ESEVISREFAGKAGV-----ASVDPSTIKQVDYAG-SFG-DAAKGKTIAAQVAAGAD- 223
OY 260 PSTLLAVAGPL-TEIFSDIIANQNR-----YLIGVDTQSL-VY---KTKNKF 306
DB 224 -TVYVAGGTGAGVFAAKSLNESRPENKVVYGVDDQDEAGKYTSKDKGSNFV 281
OY 307 SILKMGYSVSVLSDLYTKKSNRNLAGFEFGKKSATVYLGIKDRFVDIADTSLGNDK 366
DB 282 STLKQVGTIVDI-----SNKAREGEFPGQ---YIVSLDKGVADIATNLSEBK 330
OY 367 KIATFAISEAKKEFEKTKTTP 388
DB 331 K---AVEDAKAKIIDGSVKVP 348

RESULT 4
C97965

conserved hypothetical protein spr0747 [Imported] - Streptococcus pneumoniae (strain
C:Species: Streptococcus pneumoniae
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001
C:Accession: C97965
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Dehoff, B.S.
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Balz, R.H.; Jaskunas, S.
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: C97965
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-374 <KUR>
A:Cross-references: GB:AE007317; PIDN:AAK99551.1; PID:g15458340; GSPDB:GN00174
C:Genetics:
A:Gene: spr0747

Query Match 9.2%; Score 198; DB 2; Length 374;
Best Local Similarity 25.4%; Pred. No. 4.6e-05;
Matches 97; Conservative 54; Mismatches 147; Indels 84; Gaps 18;

OY 33 RKSEIMAAKADANKHFGMLMAIVTAGTVNDNSFNQSGWEALQOLGA---LTGGEITSVD 89
DB 49 RSSRNAASSDVK-----TKAALVDTGVDKSFQNSAMEGLQAMGKEHNLKDNKGFYF 104
OY 90 SSTALE-----GKYSLLANTKNVVLGSGFHGDATFRLKIPENKOLFTEKNI 140

Db 343 VFP 345

RESULT 7

ABC transporter (lipoprotein) homolog yufN - *Bacillus subtilis*
C:Species: *Bacillus subtilis*
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000

C:Accession: C70009

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C.: Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capiano, V.; Carter, N.M.; Ch
A.: Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizal, A.; Gall
A:Authors: Harwood, C.R.; Henuat, A.; Hildert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
V. M.; Ogawa, K.; Ogilawa, A.; Oudaga, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portelle
A:Authors: Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon
A:Authors: Schleich, S.; Schuster, R.; Scofield, F.; Sekiguchi, J.; Sekowska, A.; Serot
A:Authors: M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winkler, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A:Reference number: A6580; MUID:98044033

A:Accession: C70009

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-350 <KUN>

A:Cross-references: GB:299120; GB:AL009126; NID:92655613; PIDN:CAB15143.1; PID:92655650

A:Experimental source: strain 168

C:Genetics:

A:Gene: yufN

C:Superfamily: ABC transporter yufN

Query Match

Best Local Similarity 8.4%; Score 180.5; DB 1; Length 350;
Matches 101; Conservative 60; Mismatches 144; Indels 85; Gaps 20;

37 IMAAANKKHKFG-----LMAIVTAGVNDNSFNQSGWEAIOQALG-----G 83

Db 10 ILGACGSEKSGSEKSKKRVANVTGVDGSDKSPNOSANEGLQAKGKELKKKNG 69

OY 84 ITVSSTAELEGRYSLSLANTKNVWVLSGFQHDAPFLMKLIPENKOLFTEKNIIIG 143

OY 70 YDLQSKSDADYTNMLKLARENFDLIYGVYLMEDSIS---EIAQR-----KMTNFAI 121

OY 144 ID-WDTENVIPTRKYNILTYTEAGMLAGYANNSFLAKFPSPPTKRSAT-VIGGGIS 201

Db 122 IDAVVDKDNV-----ASITFEQEGSFLVGYAAL-----SSKSGKIGFVGCMES 166

OY 202 PAVTFDIFAGYLAGIKAMWLKNSDKTKITTDKIEINLGFVDVT--STKERLEQJASKDK 259

Db 167 ELIKFEVEYFRAGVOAVNPK-----AVVEVKYAGGFDRKADVKATRESM---YKSG 214

OY 260 PSTLLAVAGPL-TEIFSDIIAN-----QNDRYLIGVDTDQ---SLVYTKTKNKEFTSL 309

Db 215 VDVYHSAGAGTGVETE-AKMLKEDPKRDVWVIGVDKDYAEQVGSTGNDNVLTSMV 273

OY 310 KNLGVSVSVDLYTKKSNRNLAGFEFGKSAIVYLGIDRFVDIADTSLGENDKILA 369

Db 274 K-----KVTVEDV--TKRASDCKPFGG-----TLTYGLDQDGVGISPKONLSDDVI- 321

OY 370 TEAISEAKKEF-----EKKTKTPAE 390

Db 322 -KAVDKKKKIIDGLEIPATEKELETKFAE 350

RESULT 8

basic membrane protein - *Thermotoga maritima* (strain MSB8)
C:Species: *Thermotoga maritima*

C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000

C:Accession: F72418

R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hic
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,
C.M.

Nature 399, 323-329, 1999

A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome
A:Reference number: A7200; MUID:99287316

A:Accession: F72418

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-359 <ARN>

A:Cross-references: GB:AE001696; GB:AE000512; NID:94980582; PIDN:ADJ5196.1; PID:9498

A:Experimental source: strain MSB8

C:Genetics:

A:Gene: TM0102

C:Superfamily: ABC transporter yufN

Query Match 8.3%; Score 178.5; DB 2; Length 359;
Best Local Similarity 23.9%; Pred. No. 0.00071;
Matches 94; Conservative 59; Mismatches 177; Indels 63; Gaps 16;

OY 33 RSEFMKADANKKFGMLMAIVTAGVNDNSFNQSGWEAIOQALGALTGGETSVDS- S 91

Db 2 RFEIVISLMIFVALFGFKVIVTVGGLGDKSFNDGTWAGIKQAEELGIEAKVIOSE 61

OY 92 TAELEGRYSLSLANTKNVWVLSGFQHDAPFLMKLIPENKOLFTEKNIIIGDWTDE- 150

Db 62 QSDYIPNLSKAEEADLVFVAVFMNTNDFVAKQYPT-----YVGIDITTPEG 112

OY 151 NVIPTRKYNILTYTEAGMLAGYANNSFLAKFPSPPTKRSATVIGGIS-PATDEFA 209

Db 113 QILP--NVLTFFKQEAFLVGYAAM-----TKGMVGFQIPPIPERRY 161

OY 210 GYLACIKAMKNSDKTKI-----TTDKIEINLGFVDVTSTKLEOI----- 254

Db 162 GEAADIKTYVLHK-KNKKILNGYTDPEPRKGLDMSQFAEGADYVFNAGCAGNGV 220

OY 255 --ASKDKPSTLLAVAGPLTEIFSDIIANONDRYLVGVDQSLVYTKKNKFPSTILNL 312

Db 221 IENAEKESA-LAGSDKLVLDIYTTNGKGFAGVMDQDY---APGAVLSAMKRV 276

OY 313 GYSVESVSLDYTKKSNRNLAGFEFGKSAIVYLGIDRFVDIADTSLGENDKILATEA 372

Db 277 DVASYGVWVAYE-----CTFEGRH-----VLGISDAVGI--SPMKYTKGLVPRV 322

OY 373 ISE---AKKEPEKTKTP-AEVKTKLEIPEM 401

Db 323 IAEILYLERLKMKEGTLKVPETQEEIDAEVPOI 355

RESULT 9

basic membrane lipoprotein B (bmbp) - *Lyme disease spirochete*C:Species: *Borrelia burgdorferi* (*Lyme disease spirochete*)

C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 29-Sep-1999

C:Accession: E70147; I40290; I40242

R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh
son, D.; Peterson, J.; Kertavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Yu
Boman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.

Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venter, J.C.

A:Title: Genomic sequence of a Lyme disease spirochete, *Borrelia burgdorferi*.

A:Reference number: A70100; MUID:98065943

A:Accession: E70147

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-341 <KLE>

A:Cross-references: GB:AE001143; GB:AE000783; NID:92688279; PIDN:AC66758.1; PID:9268

A:Experimental source: strain B31

R:Simpson, W.J.; Clepax, W.
FEMS Microbiol. Lett. 119, 381-388, 1994

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2002, 15:30:19 ; Search time 17.51 Seconds

(without alignments)
935.372 Million cell updates/sec

Title: US-09-676-249a-4

Perfect score: 2155

Sequence: 1 MMDKRTKREKESADNKNQI.....KQPDKQI:SDDKITDINNL 423

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	183	8.5	325 1	BMPA_BORGA
2	181	8.4	353 1	THRC_TREPA
3	180.5	8.4	350 1	YOFN_BACSU
4	178	8.3	341 1	BMPB_BORBU
5	173	8.0	357 1	TCSA_LISIN
6	169	7.8	341 1	BMPB_BORGA
7	168	7.8	357 1	TCSA_LISMO
8	166	7.7	339 1	BMPA_BORAF
9	165.5	7.7	341 1	BMPD_BORBU
10	161	7.5	339 1	BMPA_BORAF
11	153	7.1	341 1	BMPB_BORAF
12	152.5	7.1	591 1	Y040_MYCGE
13	146	6.8	353 1	BMPB_BORBU
14	136	6.3	547 1	MEBA_STRAU
15	129	6.0	498 1	PEOB_STRPY
16	128	5.9	657 1	Y040_MYCPN
17	124.5	5.8	981 1	SCA4_RICFE
18	123.5	5.7	626 1	CEJA_BCOLI
19	118.5	5.5	1140 1	YD76_MYCPN
20	117	5.4	798 1	YCOO_MYCPN
21	114.5	5.3	848 1	Y047_SCHPO
22	114	5.3	1290 1	XCPG_XENLA
23	114	5.3	626 1	PARC_BORBU
24	114	5.3	943 1	LBPA_NEIMB
25	113.5	5.3	552 1	FLCK_SALTY
26	113.5	5.3	998 1	SCA4_RICAK
27	112	5.2	550 1	FLIC_SHIFL
28	111.5	5.2	1111 1	KIP1_YEAST
29	110.5	5.1	705 1	PCGE_FLAME
30	110.5	5.1	1164 1	BAG_STRAG
31	109.5	5.1	672 1	Y042_MYCPN
32	109	5.0	445 1	MAPA_STRMU
33	108.5	5.0	626 1	CEJB_BCOLI

34	108.5	5.0	944 1	LBPA_NEIMA	O91k4 neisseria m
35	108	5.0	3418 1	BRC2_HUMAN	P51587 homo sapien
36	107.5	5.0	794 1	YC84_MYCPN	P75493 mycoplasma
37	107.5	5.0	2022 1	ANTI_ONCVO	P21249 onchocerca
38	107	5.0	770 1	ECEL_HUMAN	P42892 homo sapien
39	107	5.0	1287 1	RPOI_FOWPV	O91593 fowlpox vir
40	106	4.9	429 1	YN8H_YEAST	P53729 saccharomyc
41	106	4.9	624 1	HTPG_BUCAI	P57555 buchnera ap
42	106	4.9	1076 1	RPOB_ASTLO	P27059 astasia lon
43	105.5	4.9	1176 1	SLAP_BACSH	P38537 bacillus sp
44	105.5	4.9	1178 1	YN17_YEAST	P48231 saccharomyc
45	105	4.9	317 1	PLC_LISMO	P34024 listeria mo

ALIGNMENTS

RESULT ID	1	BMPA_BORGA	STANDARD;	PRT;	325 AA.
AC	O31357; O31360;				
DT	15-DEC-1998 (Rel. 37, Last sequence update)				
DT	15-DEC-1998 (Rel. 37, Last annotation update)				
DE	Basic membrane protein A precursor (Immunodominant antigen P39) (Fragment).				
GN	BMPA.				
OS	Borrelia garinii.				
OC	Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.				
OX	NCBI_Taxid=295119;				
RN	(1)				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-PBI, AND PLI;				
RX	MEDLINE=98010210; PubMed=9350727;				
RA	Roesler D., Hauser U., Wilske B.;				
RT	"Heterogeneity of BmpA (P39) among European isolates of Borrelia burgdorferi sensu lato and influence of interspecies variability on serodiagnosis."				
RL	J. Clin. Microbiol. 35:2752-2758(1997).				
CC	-1- FUNCTION: NOT KNOWN: IMMUNOGENIC PROTEIN.				
CC	-1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor (Probable).				
CC	-1- SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).				
DR	EMBL: X87244; CAA5883.1; -				
DR	EMBL: X97238; CAA5877.1; -				
DR	InterPro: IPR003760; Bmp.				
DR	InterPro: IPR000437; Prok_lipoprot.				
DR	Pfam: PF02608; Bmp; 1.				
DR	PROSITE: PS00013; PROKAR_LIPOPROTEIN; PARTIAL.				
KW	Antigen; Membrane; Lipoprotein; Signal.				
FT	NON_TER	1			
FT	SIGNAL	<1	3		PROBABLE.
FT	CHAIN	4	325		BASIC MEMBRANE PROTEIN A.
FT	LIPID	4	4		N-ACYL DIGLYCERIDE (PROBABLE).
FT	VARIANT	8	8		G -> D (IN STRAIN PLI).
FT	VARIANT	12	12		S -> N (IN STRAIN PLI).
FT	VARIANT	93	93		P -> S (IN STRAIN PLI).
FT	VARIANT	111	111		T -> A (IN STRAIN PLI).
FT	VARIANT	180	180		N -> D (IN STRAIN PLI).
FT	VARIANT	205	205		S -> G (IN STRAIN PLI).
FT	VARIANT	239	239		I -> V (IN STRAIN PLI).
FT	VARIANT	243	243		A -> S (IN STRAIN PLI).
FT	VARIANT	250	250		S -> A (IN STRAIN PLI).
FT	VARIANT	257	257		N -> S (IN STRAIN PLI).

RP SEQUENCE FROM N.A.
 RC STRAIN-168;
 RA Oudega B., Koningssteijn G., Duesterhoeft A., Hilbert H.;
 RL Submitted (APR-1997) to the EMBL/GenBank/DBA databases.
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 CC (Potential).
 CC -1- SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: 293937; CAB07936.1; -
 DR EMBL: 299120; CAB5143.1; -
 DR Subtilist; BG12349; yuFn.
 DR InterPro: IPR003760; Bmp.
 DR Pfam: PF02608; Bmp; 1.
 DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 DR Hypothetical protein; Membrane; Lipoprotein; Signal;
 DR Complete proteome.
 FT SIGNAL 1 13 PROBABLE.
 FT CHAIN 14 350 HYPOTHETICAL LIPOPROTEIN YUFN.
 FT LIPID 14 14 N-ACYL DIGLYCERIDE (PROBABLE).
 SQ SEQUENCE 350 AA; 37349 MW; 16D5176A52A99284 CRC64;

Query Match 8.4%; Score 180.5; DB 1; Length 350;
 Best Local Similarity 25.9%; Pred. No. 0.00027;
 Matches 101; Conservative 60; Mismatches 144; Indels 85; Gaps 20;

QY 37 IMAAKADANKHFG-----LNMAITYAGCTVANDSFNOSGWEAIOQLALTC-----G 83
 DB 10 ILGACNSEKSSGSGCKNKFSYAMVTDGCVDDKSFNSAMWGICQAFCKENILKGRNG 69
 QY 84 EITSVDSSTAELEGKSSSLANTKNWVLSGFQGDFAFTWMLKIPERKQLTENITLG 143
 DB 70 YDVLQSKSDADYTNLNLKLAARENFDLIYGVYLMEDSIS---EIAADR-----KNTFAI 121
 QY 144 ID-WTDENVIPYGRYINLTYTEAGWLAGYNASFLAKFPSPDKNSAI-VIGGIS 201
 DB 122 IDAVVVKDNY-----ASTFKEQEGSFIVGVAAL-----SKSKGKIGFVGGMES 166
 QY 202 PATYDPIAGLAGIKAMNKNKSKRTKITDKIENLGFVQDT--STKERLQIASKDK 259
 DB 167 ELIKKEEVRAGVQAVNPK-----AVEEVKYGGEFKADVGRKVTAFESM---YKSG 214
 QY 260 PSTILAVAGPL-TEIFSDITAN-----QNDRYLIGVDTDQ--SLVYTKRKKFTSIL 309
 DB 215 VDVYHSAGATGTVFTE-AKNLKREDPRKDVAVIGVDKQVGEQVGGEDDVTILTSV 273
 QY 310 KNLGYVSFVSLDYTKRKSNNLAGFEKGSATVYLGIKDRPVDIADTSLSGNDKKLA 369
 DB 274 K-----KVDYVEVDY-TKKASDGKFPGE-----TLTYGLDDQGVGISPSKQMLSDVI- 321
 QY 370 TEAISEAKKEF-----EKKKTIPAE 390
 DB 322 -KAVDKMKKKTIIDGLEIPATEKELTKFKAE 350

RESULT 4
 BMBP BORBU STANDARD: PR: 341 AA.
 AC Q45011; Q07954; Q31317; Q50168; Q44858;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 GN Basic membrane protein B precursor.
 GE BMBP OR BR0382.
 OS Borrelia burgdorferi (Lyme disease spirochete).

OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SH-2-82;
 RX MEDLINE=94327086; PubMed=8050720;
 RA Simpson W.J., Cieplak W., Schrupf M.E., Barbour A.G., Schwan T.G.;
 RT Nucleotide sequence and analysis of the gene in Borrelia burgdorferi
 RL encoding the immunogenic P39 antigen";
 RL FEMS Microbiol. Lett. 119:381-388(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-297;
 RX MEDLINE=97132632; PubMed=8978084;
 RA Aron L., Toth C., Godfrey H.P., Cabello F.C.;
 RT Identification and mapping of a chromosomal gene cluster of Borrelia
 RL burgdorferi containing genes expressed in vivo.";
 RL FEMS Microbiol. Lett. 145:309-314(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 35210 / B31;
 RX MEDLINE=98010210; PubMed=9350727;
 RA Roessler D., Hauser U., Wilske B.;
 RT Heterogeneity of BmpA (P39) among European isolates of Borrelia
 RL burgdorferi sensu lato and influence of interspecies variability on
 RL serodiagnosis.";
 RL J. Clin. Microbiol. 35:2752-2758(1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 35210 / B31;
 RX MEDLINE=9805943; PubMed=9403685;
 RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
 RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.R., Gwinn M.,
 RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D., Hanson M.,
 RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
 RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
 RA Uitterlidge T., Wathey L., McDonald L., Artiach P., Bowman C.,
 RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
 RA Smith H.O., Venter J.C.;
 RT "Genomic sequence of a Lyme disease spirochete, Borrelia
 RL burgdorferi";
 RL Nature 390:580-586(1997).
 RL [5]
 RN SEQUENCE OF 1-179 FROM N.A.
 RP STRAIN-212;
 RX MEDLINE=95111614; PubMed=7812434;
 RA Ojalim C., Davidson B.E., Saint-Girons I., Old I.G.;
 RT Conservation of gene arrangement and an unusual organization of rRNA
 RT genes in the linear chromosomes of the Lyme disease spirochaetes
 RL Borrelia burgdorferi, B. garinii and B. afzelii.";
 RL Microbiology 140:2931-2940(1994).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 CC (Probable).
 CC -1- SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L24194; AAA72407.1; -
 DR EMBL: U49938; AAC44713.1; -
 DR EMBL: X81517; CA57227.1; -
 DR EMBL: AE001143; AAC66758.1; -
 DR EMBL: L35050; AAC41402.1; -
 DR TIGR: BB0382;
 DR InterPro: IPR003760; Bmp.
 DR Pfam: PF02608; Bmp; 1.
 DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 DR Membrane; Lipoprotein; Signal; Complete proteome.
 KW

FT SIGNAL 1 14 PROBABLE.
 FT CHAIN 15 341 BASIC MEMBRANE PROTEIN B.
 FT LIPID 15 341 N-ACYL DIGLYCERIDE (PROBABLE).
 FT VARIANT 45 45 S -> A (IN STRAIN 212).
 FT VARIANT 233 233 A -> T (IN STRAIN B31).
 FT VARIANT 318 318 V -> I (IN STRAIN B31).
 FT CONFLICT 53 53 R -> P (IN REF. 3).
 FT CONFLICT 67 67 S -> P (IN REF. 3).
 SQ SEQUENCE 341 AA; 37549 MW; C26A0D4B1D52P39F CRC64;

Query Match 8.3%; Score 178; DB 1; Length 341;
 Best Local Similarity 26.9%; Pred. No. 0.0037;
 Matches 101; Conservative 58; Mismatches 124; Indels 92; Gaps 25;

OY 44 ANKHGFLNMAIYAGCTVNDNSFNOSGWEAIOALGTCGETSVDSVAELEGKYS-- 101
 DB 24 SSKTKIRISMLV---DGVLDKSPFNSSANALRLKDKPFENIEVFS--CAISGVYSSIV 78
 DB 102 --LAINTNKN---VWVLSGFHGDADFRLWKIPENKQLETERNII--ILGIDMTDTENVI 153
 DB 79 SLDNLKRRGSDLIW-LVGYMLTDA--SLVASENPKI--SYGIIDPIYGDVQIPEMLI 133
 OY 154 PGRGRIINTYKTEEGAGMLAGY--ANASFLAKKFPSPDPTKRSALVIG--GGISPAVTD-FI 208
 DB 134 A-----VVERVEOGAFGLAGYIAAKKSFSGK-----IGFIGMGKNTVDAER 174
 OY 209 AGYLAGIKAMNLKNSDKRKITTDKRIEINLGFVDVDTSTKERLEQIASK--DKPSTLLAV 266
 DB 175 YYESGAKKAN-KDIEIIESEYNSFSVDIG-----RTIASMYSKGIIDVIHF 221
 OY 267 AGPLREI-FSDIIANQND-RYLIGVTDQSLVYTRKNEFTSILKNLGSVFSVLSLDY 324
 DB 222 AAGLAGIGVIEAKNLDGCVYVIGADQDS--YLAPKN-FITSVKNIGIDALYLITGE-Y 277
 OY 325 TKKSNSRLNAGEFEFGKSKATYVVLGIDKRVVDIADTSLEGNCKKLATEAISEAKKEEET 384
 DB 278 INNNAV-----MEGK-----VQMGRLDVGIGLPMAN-----EF-EYI 310
 OY 385 KTIPEAEVKTLEIP 399
 DB 311 KYLERKIVKKEIIVP 325

RESULT 5
 TCSA_LISIN STANDARD; PRT: 357 AA.
 092BM7:
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE CD4+ T cell-stimulating antigen precursor.
 GN TCSA OR LIN1425.
 OS Listeria innocua.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Listeria.
 OX NCBI_TaxID=1642;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CLIP 11262 / Serovar 6a;
 RX MEDLINE=21537279; PubMed=11679669;
 RA Glaser P., Frangeul L., Buchrieser C., Rusnlok C., Amend A.,
 Baquero F., Berche P., Blocher H., Brandt P., Chakraborty T.,
 Charbit A., Chetoui F., Couve E., de Daruvar A., Dehoux P.,
 Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
 Entian K.-D., Fehl H., Garcia-del Portillo F., Garrido P.,
 Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
 Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkát G.,
 Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
 Nordstedt G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
 Rammel B., Rose M., Schluter T., Simoes N., Tierrez A.,
 Vaquer-Boland J.-A., Voss H., Weiland J., Cossart P.,
 RT Comparative genomics of Listeria species.;

RL Science 294:849-852(2001).
 CC -I- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 CC (Probable).
 CC -I- SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-stb.ch/announce/>
 CC or send an email to license@isb-stb.ch).
 CC -----
 DR EMBL: AL596168; CAC96656.1; -
 DR Listlist: LIN01425; -
 DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 DR Antigen: Membrane; Lipoprotein; Signal; Complete proteome.
 FT SIGNAL 1 22 PROBABLE.
 FT CHAIN 23 357 CD4+ T CELL-STIMULATING ANTIGEN.
 FT LIPID 23 23 N-ACYL DIGLYCERIDE (PROBABLE).
 SQ SEQUENCE 357 AA; 38357 MW; 2EDA2AA9820A3FBC CRC64;

Query Match 8.0%; Score 173; DB 1; Length 357;
 Best Local Similarity 24.2%; Pred. No. 0.00079;
 Matches 101; Conservative 61; Mismatches 138; Indels 118; Gaps 23;

OY 27 SGLV-----NEKSEIMAAKADANKHFLNMAIYAGGVNDNSFNOSGWEAIOQLG 78
 DB 16 SGVVLGACGSSDDDKS-----GDCKSKDFTVAMVTDGTGVDNRSNOSAMEGLQKFG 69
 OY 79 ALT-----GGEITSVDSSTAELEGKYSILANTKNVWLS-----GFOHGDADFRLWK 126
 DB 70 KANDEKMGDGVNLYQSASEADYK-----ITLNAVAVSDDLIGIGYKLLKDAIEVSK 123
 OY 127 -IPEKQLETERNIIILGIDMTDTENVIPTGRYINLTYKTEEGAGLAGYANASFLAKFP 185
 DB 124 OKPKNOFAIVDDTI-----DDRNVVYSIG-----FRDNGSYLVGVAGL----- 163
 OY 166 SDPTKRSALVIGGISPAVTD-FIAGYLAGIKAMNLKNSDKKKTITTDKRIEINLGFVDO- 243
 DB 164 -TTTKNKVGYGVKAGVIDFEAGFTAGVA-----VNPNQIDVQY 205
 OY 244 --DTSTKERLEQIASKDKPS---TLVAGPLTEIFSDIIANQND-----YLLIGVDT 291
 DB 206 ANDPAKADKGGQIASSMVSQGVDFEHAAGTGNGVFAE-AKNLKKKPSRAVWVIGDR 264
 OY 292 DO-----SLVYTKTKNKFSTILKNLGSVFSVLSLDYTKKSNSRLNAG-FEFGKKSAT 344
 DB 265 DQWDEGKYATANDKDYNTLTSEIRVDIAV---DDLAT-----RTKAGDFPGGTFK--- 312
 OY 345 VYLGIKDFVADIADTSLEGNCKKLATEAISEAKKEFEKTTIPEAEVKTLEIPEMP 402
 DB 313 IYEGDKDAVGLSE-----HODNISKDVLAKE---ETKQKIVDGD-----IKVEKP 357

RESULT 6
 BMPB_BORGA STANDARD; PRT: 341 AA.
 ID BMPB_BORGA
 AC O31362;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Basic membrane protein B precursor.
 GN BMPB.
 OS Borrelia garinii.
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=29519;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PRT;
 RX MEDLINE=96010210; PubMed=9350727;
 RT Roesler D., Hauser U., Wilske B.;

*Heterogeneity of Bmpa (P39) among European isolates of *Borrelia burgdorferi sensu lato* and influence of interspecies variability on serodiagnosis." ;
 RL J. Clin. Microbiol. 35:2752-2758(1997).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor (Probable).
 CC -1- SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 DR EMBL: X81518; CA57238.1; -
 DR InterPro: IPR003760; Bmp.
 DR Pfam: PF02608; Bmp; 1.
 DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 KW Membrane; Lipoprotein; Signal.
 FT SIGNAL 1 14 PROBABLE.
 CHAIN 15 341 BASIC MEMBRANE PROTEIN B.
 LIPID 15 15 N-ACYL DIGLYCERIDE (PROBABLE).
 SQ SEQUENCE 341 AA; 37236 MW; 442BEF0BEDDC9A CRC64;
 Query Match 7.8%; Score 169; DB 1; Length 341;
 Best Local Similarity 26.8%; Pred. No. 0.0013;
 Matches 91; Conservative 55; Mismatches 117; Indels 76; Gaps 20;
 QY 59 GTVNDNSFNOSGWEAIQOLGALTGETSVDSSTAELEGKYS--LANTNRN---VW 110
 DB 36 GVLDDKSFNSARALLRLREDFPENIEKVFSSA--SGVYSVSDLDLKNKNGSDLIW 93
 QY 111 VLSGFGHGAFTWMLKPEKOLETEKNITILG-IDMTDENYIPIRTYINLYTKTEAG 169
 DB 94 -LVGYMLTDA-----SLVSLNPKLISYGIIDPVSDVDQIPKNLGVVERIIOGA 143
 QY 170 WLAGYANASFLAKFPSPDPTKRSATVIGGISPAVTD-FIAGYLAGIKAMNLSKDKTK 228
 DB 144 FLAGY-----IAAKKVS-----GKIFIGVKGDIYDAFRYEGAG-----AKADGIE 189
 QY 229 ITTDR---TEINLGFVDPTSTKERLEQIASK--DKPSTLLAVAGHLEIFSDIANON 282
 DB 190 IYSEYSNSFSDVNI-----RAYANKVYAKGIDIIHPAAGLAGI--GVIEAPK 235
 QY 283 D-----KYLICVDPTDQSLVYTKTKNKFPTSLKNGSVESVLSDLTYKKNSRNLAGFER 338
 DB 236 ELGDGYVIGADDDSHL---APRNFPTSVIKKVGALYLTSE-VYKNNNT-----WEG 286
 QY 339 GKKSATVYLGIKDFVDIADTS---LEGNDKILATEAI 373
 DB 287 GK---IIQGLRDGVVGLSNANKFEYIKVIERKIVNEEI 322
 RESULT 7
 TCSEA.LISMO STANDARD; PRT; 357 AA.
 AC 048754;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE CD4+ T cell-stimulating antigen precursor.
 GN TCSEA OR LMO1388.
 OS *Listeria monocytogenes*.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; *Listeria*.
 OX NCBI_TaxID=1639;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BGD-e / Serovar 1/2a;
 RX MEDLINE=21537279; PubMed=11679669;

RA Glaser P., Frangoul L., Buchrieser C., Rusniok C., Amend A.,
 RA Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
 RA Charbit A., Cheretouani F., Couve E., de Daruvar A., Deboux P.,
 RA Dommien E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
 RA Ertan K.-D., Fsih H., Garcia-del Portillo F., Garrido P.,
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Haut J., Jackson D.,
 RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapat G.,
 RA Meduno E., Maitounan A., Mata Vicente J., Ng E., Nedjari H.,
 RA Nordiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
 RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
 RA Vazquez-Boland J.-A., Voss H., Wenland J., Cossart P.,
 RT Comparative genomics of *Listeria* species." ;
 RL Science 294:849-852(2001).
 RN [2]
 RP SEQUENCE OF 1-252 FROM N.A.
 RC MEDLINE=85EO-1167;
 RX MEDLINE=96096448; PubMed=7500019;
 RA Sanderson S., Campbell D.J., Shastri N.;
 RT Identification of a CD4+ T cell-stimulating antigen of pathogenic
 RT bacteria by expression cloning." ;
 RL J. Exp. Med. 182:1751-1757(1995).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor (Probable).
 CC -1- SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 DR EMBL: AL591979; CAC99466.1; -
 DR EMBL: S80336; AAB35725.2; ALT-TERM.
 DR LISTPRO: LMO01388; -
 DR InterPro: IPR003760; Bmp.
 DR Pfam: PF02608; Bmp; 1.
 DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 KW Antigen; Membrane; Lipoprotein; Signal; Complete proteome.
 FT SIGNAL 1 22 PROBABLE.
 CHAIN 23 357 CD4+ T CELL-STIMULATING ANTIGEN.
 LIPID 23 23 N-ACYL DIGLYCERIDE (PROBABLE).
 SQ SEQUENCE 357 AA; 38415 MW; 83605B8B6419C8D1 CRC64;
 Query Match 7.8%; Score 168; DB 1; Length 357;
 Best Local Similarity 23.8%; Pred. No. 0.0016;
 Matches 98; Conservative 62; Mismatches 145; Indels 106; Gaps 22;
 QY 26 ISGLVNERKSEIMAKADANKHFLMAIYTAGTVNDNSFNOSGWEAIQOLGALT---- 81
 DB 19 ILGACSSSDPKSSDKSSKDF--TYAMVTIDIGVDRSFNSAMGLOKFGKANDMEK 76
 QY 82 GELTIVSDSTAELEGKYSLSANTNKNVWVLS-----GFQHGDAFTWMLK-IPENKQ 132
 DB 77 GTDGYNYLOSSEADYK-----TNLNTAVRSDYDLTYGKYLKDAIEEYSKQPRNOF 130
 QY 133 LFTFKNIIILGIDMTDENYIPIGRYINLYTKTEAGWLAGYANASFLAKFPSPDKRS 192
 DB 131 AIVDDTI-----DDRDNVVSTIG-----FKNDGSGYLVGVAGL-----TTKTN 168
 QY 193 AIVGGISPAVTD-FIAGYLAGIKAMNLSKDKTKITTDKIEINLGFVQ--DPTSK 248
 DB 169 KVGFGVKGKVIDRFAGGFTAGYKA-----VAPNAQIDVQVANDPRAKA 212
 QY 249 ERLEQIASDKPS---TLAVAGHLEIFSDIANONDR-----YLIGVTDQ----- 293
 DB 213 DKGOIASSMYSGVDYIFHAAGGTGNGVFAE-AKNLKKDPSRAVWVIGVDRDQWDEK 271
 QY 294 -SLVYTKTKNKFPTSLKNGSVESVLSDLTYKKNSRNLAG-FEEGKKSATVYLGIKD 351
 DB 272 VTANDKDVNVTLTSEIKRDIADV---EDLAT-----RAKAGDFPGTK--TEYGLDK 319

```

OY 352 RYVDIADTSLKGNKKLATEAISEAKKEFEKTKTIPAEVRKLTIEPMP 402
DB 320 DAVGLSE-----HODNISKVDLAKVE---EYKQKIVDGD-----IKVPEKP 357

RESULT 8
BMPA_BORAF STANDARD: PRT: 339 AA.
AC O31280: O31281: O31282: O31283:
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Basic membrane protein A precursor (Immunodominant antigen P39).
GN BMPA.
OS Borrelia afzelii.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OY NCBI_TaxID=29518;

[1]
SEQUENCE FROM N.A.
RC STRAIN-PKO, PWUDI, PLE, AND PLJ7;
RX MEDLINE=98010210; PubMed=9350727;
RA Roessler D., Hauser U., Wilske B.;
RT "Heterogeneity of BmpA (P39) among European isolates of Borrelia burgdorferi sensu lato and influence of interspecies variability on serodiagnostics."
RL J. Clin. Microbiol. 35:2752-2758(1997).
CC -1- FUNCTION: NOT KNOWN; IMMUNOGENIC PROTEIN.
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor (Probable).
CC -1- SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X81516; CAA57236.1; -
DR EMBL: X97237; CAA65876.1; -
DR EMBL: X97239; CAA65878.1; -
DR EMBL: X97241; CAA65880.1; -
DR InterPro: IPR003760; Bmp.
DR Pfam: PF02608; Bmp. 1.
DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
CC Antigen; Membrane; Lipoprotein; Signal.
FT SIGNAL 1 17 PROBABLE.
FT CHAIN 18 339 BASIC MEMBRANE PROTEIN A.
FT LIPID 18 18 N-ACYL DIGLYCERIDE (PROBABLE).
FT VARIANT 125 125 A -> S (IN STRAIN PLE).
FT VARIANT 214 214 I -> T (IN STRAIN PLJ7).
FT VARIANT 229 229 A -> P (IN STRAIN PKO).
FT VARIANT 254 254 I -> V (IN STRAIN PLJ7).
FT VARIANT 268 268 L -> F (IN STRAINS PLE AND PKO).
SQ SEQUENCE 339 AA: 36966 MW: 15EADBF8CBA06F84 CRC64:

Query Match 7.7%: Score 166; DB 1; Length 339;
Best Local Similarity 23.3%; Pred. No. 0.002;
Matches 89; Conservative 68; Mismatches 143; Indels 82; Gaps 19;
OY 49 GLNMAIVTAGTVNDNSFNQSGWEAIQQLGALTGGEITSVDST-----AELEGKYSILAN 104
DB 27 GIPKVSIVTIGTFDDKSFNSALNGVKLKEFEIEVLKSSSTSYSLDLEG-----LKD 82

OY 105 TKNKNWVSGVOHQDAFTFRMLKIPFNKQLFTEKNIIILGIMDTDTENVIPGRVINTLYK 164
DB 83 AGSNILWIGIKFSDV-AKAVSLQNSEKKYAI-----IDPVYSNEPIP-ANLWGKTFR 133

OY 165 TEEAGMLAGYANASFLAKKPPSDPKRSAYIYGGISPAVMD-FLAGLAGIKAMNKLNS 223

```

```

DB 134 AOEAGFLGIYIAKV-----SKTGKIGFLGIEGDIYDAFRYGYEAGAKYAN---- 180
OY 224 DKRTITTDKI-----EINLGEVDQDSTKERLEQJASRKDPSTLLAVGPLEIFSDIIA 279
DB 181 -KDIKIFSGYIGSFSDLENG---RSVATKMYSDG-----DIHHAAGLGICAIIEVA 229

OY 280 NO-NDRYLIGVDTPOSLVYTKTKNKKFFTSILKNGYVFSVLSLYTKKSNRLAAGE 337
DB 230 KELGSGHYIIGVDEQSYL---APNNVITSTYTKDGRSL-NLLTSNYLKTWT-----FE 279

OY 338 FKKRSATYVLGIKDRFVIAOTSLKGNKKLATEAISEAKKEFEKTKTIPAEVRKLTLE 397
DB 280 GGR---LNYGLKEGVGVFV-----RNPMPKEFV-----EKEIDSLSKTIINKEVI 323

OY 398 IPEMPDKOPDKQESLDKLTID 419
DB 324 VP-----YKKESEYKFLKE 337

RESULT 9
BMPD_BORBU STANDARD: PRT: 341 AA.
AC 044743:
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Basic membrane protein D precursor.
GN BMPD OR BB0385.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OY NCBI_TaxID=139;

[1]
SEQUENCE FROM N.A.
RC STRAIN-JDI;
RX MEDLINE=96178617; PubMed=8606088;
RA Ramamoorthy R., Rovinelli L., Philipp M.T.;
RT "Molecular characterization, genomic arrangement, and expression of bmpD, a new member of the bmp class of genes encoding membrane proteins of Borrelia burgdorferi."
RL Infect. Immun. 64:1259-1264(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lachigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Uitterback T., Fujii C., Cotton M.D., McDonald L., Artlich P., Bowman C.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi."
RL Nature 390:580-586(1997).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor (Probable).
CC -1- SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U35450; AAC43984.1; -
DR EMBL: AE001144; AAB91505.1; -
DR TIGR: BB0385; -
DR InterPro: IPR003760; Bmp.
DR Pfam: PF02608; Bmp. 1.
DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.

```


Db 326 -SNKSEYEKLE 337

RESULT 11

EMBL_BORAF STANDARD: PRT: 341 AA.

AC O31284;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Basic membrane protein B precursor.
 GN BMBP.
 OS Borrelia afzelii.
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=29518;
 [1]

PP SEQUENCE FROM N.A.

STRAIN=PKO; PubMed=9350727;
 MEDLINE=98010210;

RA Roessler D., Hauser U., Wilske B.;
 "Heterogeneity of BmpA (P39) among European isolates of Borrelia
 burgdorferi sensu lato and influence of interspecies variability on
 serodiagnostics.";

RT J. Clin. Microbiol. 35:2752-2758(1997).
 RL -1 SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 (Probable).
 CC -1 SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).

CC EMBL; X81519; CA57239.1; -.

DR InterPro: IPR003760; Bmp.

DR Pfam: PF02608; Bmp; 1.

DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.

KW Membrane; Lipoprotein; Signal.

FT SIGNAL 1 14 PROBABLE.

FT CHAIN 15 341 BASIC MEMBRANE PROTEIN B.

FT LIPID 15 15 N-ACYL DIGLYCERIDE (PROBABLE).

SEQUENCE 341 AA; 37198 MW; 188EA3E3D54ACDB1 CRC64;

Query Match 7.1%; Score 153; DB 1; Length 341;

Best Local Similarity 25.4%; Pred. No. 0.012;

Matches 90; Conservative 55; Mismatches 126; Indels 84; Gaps 19;

59 GIVNDSNFOSGWEALIOGLALTGGEITSVDSSTAELEGKSS-----LANTNKN-----VW 110

36 GVLDSKSNSSANENALLKRLKADFPENIEKVS--VSGYSSISYDLDNLKMSNGLIW 93

111 VLSGFQHD-AETRWLKIPEKQKLFTEKNII--ILGIDMTDTEENVIPGRYINLTYYKTE 167

94 -LVGYMLADVLSVLENDPE-----INVGIIDPIYGDDVQIIPKLN-----IGVVERIEQ 141

168 AGMLAGYANASFLAKFSPDPKRSAYIYGGISPAVTD-FLAGLAGIKAMN-----L 220

142 GAFLAGY-----IAAKKSVS-----SKIGFIGVGKGDIVDAFFGYEAGAKYANKGIEIYS 192

221 KNSDKKLTITTDKIEIN---LGFDVODTSTKERLEQJASKRKPSTLLAVAGPLREISD 276

193 EYSNFSIDILARVANKMYSGKIDIIHPAA--GLAGVGVIEAPKEL----- 237

277 IIANONDRYLIGVPTDQSLVYTKNKKEFTSILKMLGYSVSVSLDLYTKKSNHLAGF 336

238 ----GDYVYVIGADQDQSHLAPK---NFTSYIKNVGDALYITSE-----SLKNDNW 284

337 EFGKSAATVYLGIKRFDVADIADTSLGNDKRLATEAISEAKKEFEKTKTTPAE 391

Db 285 EGK---IVQMGLRDGVVGLSNAN-----EFEYIKDLERKINKKEIIVPCNQ 328

RESULT 12

Y040_MYCGE STANDARD: PRT: 591 AA.

AC P47286;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical lipoprotein MG040 precursor.
 GN MG040.
 OS Mycoplasma genitalium.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
 OC Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2097;
 [1]

PP SEQUENCE FROM N.A.

STRAIN=ATCC 33530 / G-37;
 MEDLINE=96026346; PubMed=7569993;

RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
 Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
 Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.L.,
 Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
 Tomb J.-F., Dougherty B.A., Bolt K.F., Hu P.-C., Lucier T.S.,
 Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
 "The minimal gene complement of Mycoplasma genitalium.";

RT Science 270:397-403(1995).
 RL [2]

SEQUENCE OF 448-517 FROM N.A.

STRAIN=ATCC 33530 / G-37;

MEDLINE=94075230; PubMed=8253680;

RA Peterson S.N., Hu P.-C., Bolt K.F., Hutchison C.A. III;

"A survey of the Mycoplasma genitalium genome by using random
 sequencing.";

RT J. Bacteriol. 175:7918-7930(1993).

RL -1 SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 (Potential).

CC -1 SIMILARITY: SOME, TO T.PALLIDIUM TWC.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).

CC EMBL; U39683; AAC71256.1; -.

DR EMBL; U02125; AAD12400.1; -.

DR TIGR; MG040; -.

DR InterPro: IPR003760; Bmp.

DR Pfam; PF02608; Bmp; 1.

DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.

KW Hypothetical protein; Lipoprotein; Membrane; signal;

FT SIGNAL 1 30 COMPLETE PROTEOME.

FT CHAIN 31 591 HYPOTHETICAL LIPOPROTEIN MG040.

FT LIPID 31 31 N-ACYL DIGLYCERIDE (POTENTIAL).

SEQUENCE 591 AA; 64019 MW; 1A8FA31BB7E0928 CRC64;

Query Match 7.1%; Score 152.5; DB 1; Length 591;

Best Local Similarity 20.8%; Pred. No. 0.078;

Matches 99; Conservative 84; Mismatches 135; Indels 159; Gaps 27;

44 ANKHGELMAIYTAG--GIVNDSNFOSGWEALIOGLALTG-ELTSSVSTAELEGK-- 98

33 ATKIFDSSVQLVSDNFSNFKSFGMSREGIRSFKKSGVLDPEADSSQLQ-EGNGL 91

99 -----YSSLANTNKNVWVLSGFOHGAFTFRWLKIPENKOLFTEKNIIIGI 144

DB 92 MKRPGTSLDRATENNINNGSDVIATGFNO-----QESLQATISDPIR--- 138
 QY 145 DWITDENVIPTG-----RYINLYTEAGWLAGYANAS 178
 DB 139 -GSDKSLAKTGFIFVDGAIEKFNKRNQVPOKSPNTINSSAFPSDDOSFLTGATAV 197
 QY 179 FL-----AKRPSDPTKRSAYIGG--GIS-PATVDFIAGYLAGIKAMN-----L 220
 DB 198 YLNLNOEYFLDKSGWSTNSSNNNELTVSGFVGIALPSTLSFLNGFRLGIAYFNEVIYKHL 257
 QY 221 KMS-DKKTITITDK-----IEINIG-----FDVOD-----TST 247
 DB 258 SDPQDSSAOVITTSKQTVKQLOVANGKIKTKMTISPGQSGEYINIQDHQSGSFSDT 317
 QY 248 KERLEQIASK--DK-PSTLLAVAGPLTE-IFSDIIANQNDRYLIGVTDQSL-----VYTK 299
 DB 318 EPRATITIANNLIDKGNVAILIPIAGPQTNLVVQIARQAHNTAVIGVDSQGLLDINIDAP 377
 QY 300 TKNKF-----PSTLKNLGYSVFVSLDYLYTKSNSRLIAGE-FOKKSATVYLG- 348
 DB 378 NMDKLMGNKKIIPFESSI-KALDVAVESILSTL--EKSSQN--GYGFGGYNN-----IGT 428
 DB 349 IKDRFVDIADT-----SLEGNDKLTATPAISEAKKEFEKTR 385
 DB 429 VKNNSVGEAGEYELIDPVFMKNITSSQAMSLASLAKNAASSSDNKKRLSEVATK 485

RESULT 13
 BMPC_BORBU STANDARD; PRT: 353 AA.
 AC 050169; 044859; (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Basic membrane protein C precursor.
 GN BmpC OR BB0384.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=297;
 RA MEDLINE=95080623; PubMed=7968902;
 RA Aron L., Aleksun M., Perlée L., Schwartz I., Godfrey H.P.,
 RA Cabello F.;
 RT "Cloning and DNA sequence analysis of bmpC, a gene encoding a
 RT potential membrane lipoprotein of Borrelia burgdorferi.";
 RT FEMS Microbiol. Lett. 123:75-82(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=297;
 RA MEDLINE=97132632; PubMed=8978084;
 RA Aron L., Toth C., Godfrey H.P., Cabello F.C.;
 RT "Identification and mapping of a chromosomal gene cluster of Borrelia
 RT burgdorferi containing genes expressed in vivo.";
 RT FEMS Microbiol. Lett. 145:309-314(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35210 / B31;
 RA MEDLINE=98065943; PubMed=9403685;
 RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
 RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
 RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
 RA Peterson J., Kariavage A.R., Quackenbush J., Salzberg S., Hanson M.,
 RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
 RA Uitterback T., Watthey L., McDonald L., Artlich P., Bowman C.,
 RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
 RA Smith H.O., Venter J.C.;
 RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
 RT burgdorferi.";
 RT Nature 390:580-586(1997).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 CC (Probable).

CC -1- SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: U49938; AAC44711.1; -.
 DR EMBL: AE001143; AAC66756.1; -.
 DR TIGR: BB0384; -.
 DR InterPro: IPR003760; Bmp.
 DR Pfam: PF02608; Bmp; 1.
 DR PROSITE: PS00013; PROKAR LIPOPROTEIN; 1.
 KW Membrane; lipoprotein; signal; complete proteome.
 FT SIGNAL 1 16 PROBABLE.
 FT CHAIN 17 353 BASIC MEMBRANE PROTEIN C.
 FT LIPID 17 17 N-ACYL DIGLYCERIDE (PROBABLE).
 FT VARIANT 179 179 V -> L (IN STRAIN 297).
 SO SEQUENCE 353 AA; 39824 MW; 2E8FF607D6CAB9B0 CRC64;

Query Match 6.8%; Score 146; DB 1; Length 353;
 Best local Similarity 23.7%; Pred. No. 0.035;
 Matches 84; Conservative 53; Mismatches 131; Indels 86; Gaps 18;

QY 55 VTAGGVYNDNSFNQSGHEAIOQLG-----ALTGEITSVDNSTAELEGKY 99
 DB 34 VLAHGSFYDKGYNOYSHDGVAKLRDNGIKLITKSLRYPDEGRRLTYDE--AMEDAY 91
 QY 100 SSLANTKNVNWVSGPHQDAFTRWLKIPEKNOLFETKNIILIGI-----DWTPTENVIP 155
 DB 92 EVQKNPLNLFWLI-GYRFSPL-----SVKLSYRPIYIIGIIDAIFYGDIQ--VPK 139
 QY 156 GRYINLYTEAGWLAGYANASFLAKRPSDPTKRSAYIG--GGISPAVYDFIAGYL 212
 DB 140 NS-LAIKFRNEEAFLAGYIAAKMSRKE-----KIGFLTGMSSHVDKFGK 187
 QY 213 AGIKAMN--LKNSDKKITTTDKIEINMGFDVODTSTKELEQIA---SKDRPSTLLAV 266
 DB 188 AGIFYANPKRLRLVSKRAPSLED-----KEKGKAMALFMYKEDYGVYFPI 232
 QY 267 AGCLTEIFSDIIANQND-FLIGVDFPDSLYVTKRKPFSTILKNLGYSVFVSLDYLT 325
 DB 233 AGITGLGVYAAKELGKYVYIGLNDQSYT---APQNVTTSTIKDKVYISISSEYT- 288
 QY 326 KKSNSRLIAGEFERGKKSATVYLGIKDRFVDA-DTSLGNDKLTATPAISEARK 378
 DB 289 ---NNRVFKG-----GIIDRGKLEGVIEIVDPVLNN--RLVDEVIDLEMK 331

RESULT 14
 MERA_STRAU STANDARD; PRT: 547 AA.
 AC P08663;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Mercuric reductase (Ec 1.16.1.1) (Hg(II) reductase).
 GN MERA.
 OS Staphylococcus aureus.
 OC plasmid p1258.
 CC Bacteria; Firmicutes; Bacillus/Clostridium group;
 CC Bacillus/Staphylococcus group; Staphylococcus.
 OX NCBI_TaxID=1280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=87260937; PubMed=3037534;
 RA Laddaga R.A., Chu L., Misra T.K., Silver S.;
 RT "Nucleotide sequence and expression of the mercurial-resistance
 RT operon from Staphylococcus aureus plasmid p1258.";

RL Proc. Acad. Sci. U.S.A. 84:5106-5110(1987)
 CC -1- FUNCTION: RESISTANCE TO HG(2+) IN BACTERIA APPEARS TO BE COVERED
 CC BY A SPECIALIZED SYSTEM WHICH INCLUDES MERCURIC REDUCTASE, MERA
 CC PROTEIN IS RESPONSIBLE FOR VOLATILIZING MERCURY AS HG(0).
 CC -1- CATALYTIC ACTIVITY: Hg + NADP(+) + H(+) = Hg(2+) + NADPH.
 CC -1- COFACTOR: FAD.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- MISCELLANEOUS: THE ACTIVE SITE IS A REDOX-ACTIVE DISULFIDE BOND.
 CC -1- SIMILARITY: BELONGS TO THE PRIDINE NUCLEOTIDE-DISULFIDE
 CC OXIDOREDUCTASES CLASS-I.
 CC -1- SIMILARITY: CONTAINS 1 HMA DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/isb-sib.ch/>)
 CC or send an email to license@isb-sib.ch).

```

DR EMBL: L29436; AAA98245.1; -.
DR PIR: E29504; E29504.
DR HSSP: P11959; 1EBD.
DR InterPro: IPR001327; FAD_pyr_redox.
DR InterPro: IPR001934; HMA.
DR InterPro: IPR000815; Hg_reductase.
DR InterPro: IPR001100; pyr_redox.
DR InterPro: IPR004099; pyr_redox_dlm.
DR Pfam: PF00403; HMA; 1.
DR Pfam: PF00070; pyr_redox; 1.
DR Pfam: PF02852; pyr_redox_dlm; 1.
DR PRINTS: PRO0368; FADPNR.
DR PRINTS: PRO0945; HGRDPTASE.
DR PRINTS: PRO0411; PNDROPTASEI.
DR PROSITE: PS01047; HMA_1; 1.
DR PROSITE: PS50846; HMA_2; 1.
DR PROSITE: PS00076; PYRIDINE_REDOX_1; 1.
DR MercuLic resistance: Oxidoreductase; Flavoprotein; FAD; NAD;
KW Mercury; Redox-active center; Metal-binding; Plasmid.
FT DOMAIN 5 68 FMA.
FT NP_BIND 87 117 FAD (ADP PART) (PROBABLE).
FT DISULFID 123 128 REDOX-ACTIVE.
FT NP_BIND 378 388 FAD (FLAVIN PART) (BY SIMILARITY).
FT METAL 544 544 HG(2+) (POTENTIAL).
FT METAL 545 545 HG(2+) (POTENTIAL).
FT SEQUENCE 547 AA; 58565 MM; BOBBA5FCFA9C049 CRC64;

```

Query Match: 6.38; Score 136; DB 1; Length 547;

[illegible]

Db	314	YQKVEQNKSKSTIITYEVNGQGEQVLEADQVLYATGRKPNFTTLWLESQGVKTKKGEVLTN	373
Qy	323	LYTKKSNSRNLAAGEFEGKSSATYV-----LGIKRDFVIADT-SLEGNDKL	368
Db	374	EYLQGSNNRIYAADVTLGPQFVYVAAYEGGIIVANNALGLAKRKIDLRFPVGVTFPNPSI	433
Qy	369	ATEALISE---AKKEFEKTKTIPEAEVRKTL	396
Db	434	ATVGLTEQQAKEKGIYKTSVLPIDAVPRRL	464

RESULT	15
PEDB_STRPY	
ID	PEDB_STRPY
AC	Q99XSL
DT	16-OCT-2001 (Rel. 40, Created)
DT	16-OCT-2001 (Rel. 40, Last sequence update)
DT	16-OCT-2001 (Rel. 40, Last annotation update)
DE	Probable dipeptidase B (EC 3.4.-.-).
GN	PEDB OR SPY2066.
OS	Streptococcus pyogenes.
OC	Bacteria: Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC	Streptococcus.
OX	NCBI_TaxID=1314;

RN SEQUENCE FROM N.A.
 RC STRAIN-SF370 / ATCC 700294 / Serotype M1;
 RX MEDLINE-21192684; PubMed-11266286;
 RA Ferretti U.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.
 RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Iln S.P.,
 RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
 RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
 RT "Complete genome sequence of an M1 strain of *Streptococcus pyogenes*."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:4658-4663(2001).
 CC -1- CATALYTIC ACTIVITY: DIPEPTIDE + H(2)O -> 2 AMINO ACID.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY U34.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaborat
 CC between the Swiss Institute of Bioinformatics and the EMBL outstat
 CC the European Bioinformatics Institute. There are no restrictions on
 CC use by non-profit institutions as long as its content is in no
 CC modified and this statement is not removed. Usage by and for comm
 CC entities requires a license agreement (See <http://www.isb-sib.ch/anno>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE006627; AK34726.1; -
 SW Hydroxylase; Dipeptidase; Complete proteome.
 NO SEQUENCE 498 AA; 55499 MW; B9334DC94C19A84D CRC64;

```

Query Match          6.0%; Score 129; DB 1; Length 498;
Best Local Similarity 20.5%; Pred. No. 0.61;
Matches 99; Conservative 65; Mismatches 175; Indels 144; Gaps

QY 4 KETTREKSAANON-----KQITDYSKISGLVNERKSEIMAAKDANKFCL 50
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 63 KDNAGEGEMKWLDSNGFEYPLBHSYRSALPDVTPKNGYDE-----AGFNE-EGV 112
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 51 NMAIVTAGGVND-----NSFNOSGWEAIIQQLGALTGE---ITSVDSSTAELE 96
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 113 SMS-ATVASANDAIORIDPYVKNGLAESMTSVILPSVTAREGVALIAKIVTERGAAL 171
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 97 GKYSLLANTNNNV---VLSFGHGFMTWMLIPENKQLFETKNIILIGIMTOTRENI 153
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 172 GNIVTLADKD-GIWMYMEILSGHOVA---IKFPDDKAVPEPNFYLGHVDFNKENT 225
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 154 PTGRINLTYTEEGWLAGYANASFLAKF--PSDPTKRSALVIGGISPATVDIAGY 211
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 226 ASEDEVAKAKAKAYTEYVDGKFH---IAKSYNPPLDANRSRFSG----- 268
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 212 LAGIRANMLKNSDKRTIT-----TDKIEINLGFVDVDTSTKER----- 250
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2002, 15:31:16 ; Search time 50.25 Seconds

(without alignments)
1456.256 Million cell updates/sec

Title: US-09-676-249A-4

Perfect score: 2155
Sequence: 1 MMDKETEKEKSADNOKOI.....KQPDKQKQSLDKLTIDINNL 423

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Minimum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPTRMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	693.5	32.2	457	2	Q48902 mycoplasma
2	552	25.6	461	16	Q980L5 mycoplasma
3	391	18.1	465	2	Q9X775 mycoplasma
4	384.5	17.8	428	2	Q52311 mycoplasma
5	382.5	17.7	428	2	Q9RGX5 mycoplasma
6	379.5	17.6	428	2	Q32417 mycoplasma
7	379	17.6	429	2	Q9RGX6 mycoplasma
8	379	17.6	429	2	Q9RGX4 mycoplasma
9	377.5	17.5	428	2	Q9RGX7 mycoplasma
10	374.5	17.4	428	2	Q9RGX3 mycoplasma
11	373.5	17.3	428	2	Q9RGX6 mycoplasma
12	228	10.6	349	2	Q9EV91 clostridium
13	220.5	10.2	350	16	Q9GPH4 streptococcus
14	203.5	9.4	350	16	Q9GPH9 streptococcus
15	202	9.4	350	16	Q97L60 streptococcus
16	192	8.9	357	16	Q97L60 clostridium

17	182	8.4	348	2	Q9AK41 streptomyc
18	178.5	8.3	359	16	Q9XWV7 thermotoga
19	174	8.1	342	16	Q9A1P7 streptococ
20	173	8.0	357	16	Q92BW7 listeria in
21	170	7.9	264	2	Q31358 borrelia ga
22	168.5	7.8	516	16	Q9PRD3 ureaplasma
23	168	7.8	341	2	Q9FED3 borrelia bu
24	164	7.6	347	2	Q9AK42 streptomyc
25	162.5	7.5	525	16	Q9PQR5 ureaplasma
26	161	7.5	339	2	Q9S6C1 borrelia bu
27	161	7.5	339	2	Q87960 borrelia bu
28	160	7.4	325	2	Q9R776 borrelia bu
29	157	7.3	339	2	Q9FED2 borrelia bu
30	156	7.2	325	2	Q9R777 borrelia bu
31	150	7.0	524	16	Q9PRC9 ureaplasma
32	148	6.9	329	16	Q9RGW4 rhizobium l
33	145	6.7	264	2	Q31359 borrelia ga
34	143.5	6.7	243	2	Q31361 borrelia ga
35	143	6.6	353	2	Q93V09 borrelia bu
36	140	6.5	379	16	Q9RSQ4 deinosoccus
37	137.5	6.4	889	2	Q9LP61 moraxella c
38	134	6.2	539	16	Q9PQ10 ureaplasma
39	133.5	6.2	1786	5	Q9UDP0 plasmidium
40	132.5	6.1	330	16	Q926D6 rhizobium m
41	132	6.1	326	16	Q92X35 rhizobium m
42	131.5	6.1	1558	5	Q96275 plasmidium
43	127	5.9	912	2	Q9XD52 moraxella c
44	127	5.9	2045	16	Q9AKK5 streptococ
45	126.5	5.9	894	2	Q9L962 moraxella c

ALIGNMENTS

RESULT 1
ID Q48902 PRELIMINARY; PRT; 457 AA.
AC Q48902;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE AG 243-5 PROTEIN PRECURSOR.
OS Mycoplasma arginini.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2094;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96163149; PubMed=8551970;
RA Ushio S., Iwaki K., Tanai M., Ohta T., Fukuda S., Sugimura K.,
RA Kurimoto M.;
RT "Metastatic promoting activity of a novel molecule, Ag 243-5 derived
RT from Mycoplasma and the determination of the complete nucleotide
RT sequence.";
RL Microbiol. Immunol. 39:393-400(1995).
DR EMBL: D16674; BAA04082.1; -
DR InterPro: IPR003760; Bmp.
DR Pfam: PF02608; Bmp; 1.
KW Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 457 AG 243-5 PROTEIN.
SQ SEQUENCE 457 AA; 50789 MW; 9AEEAB11620CB22F CRC64;

Query Match 32.2%; Score 693.5; DB 2; Length 457;
Best Local Similarity 39.0%; Pred. No. 1,9e-32;
Matches 188; Conservative 81; Mismatches 141; Indels 41; Gaps 14;

QY 5 ETTKEKSADNOKOITDYSKISGLVERSEIIMAAKADNKHFGNMAIVTAGGTYNDN 64
DB 28 ETTKEKGIIT-----RIFD-----NSFVDRQAEIEKA-----KNDPNTVLTLTAGGVODK 73
QY 65 SFNOSGWEAI---QDGLALTG--GEITSVDSSTAELGKYSILANTNKNVWVLSGFOHGD 119

```

Db 74 SFGSISWEAVLEHYDIERTKTLNDRVSOETNNQSELIGKFKFLNGKNKVMWILTGQOQO 133
QY 120 AFTMRKLIPEN-----KOLFTEKNIIILGIDW---TDENVIPTGRIYLMVTKTEBAGWL 171
Db 134 EEPKRLKOTDSNGKKYSDDLAEKKYIYAVBDLSTKEDDLKAGHIFSLILKTEBAGTI 193
QY 172 AGYANASFLAKKFPSPDPTKRSATVIGGISPAVTDFIAGYLAGIKAMNLKNSDKTKRTT 231
Db 194 AGYASSKFLAYKFPNDEAKRTIAPFGCGHAGVDFIAGFLAGIAKYNNDNPJAKVTIISD 253
QY 222 DKIEINLIGDVODTSTKERLEQIASKDKRSTLLAVAGPLTEIFSDI--ANONDRLLIIV 289
Db 254 NNINIDTGF-ISNDKTAFINGIYNKS--SLVLPVAGSLTSSVVAIKKSNKDTKYLIGV 310
QY 230 DDDQSLVYTKTKNKFPTSLIKMGYSVFSVLSDLTYTKKSNHNLAGEFGK---KSATV 345
Db 311 DTDQSKITSFA-TYFTSTSEKHLGRTIYVLDIMLKKEDSFSLGSRFKLTNRANATV 369
QY 346 YGIGDRFVDIADTSLGENDKRLATEISAEKFEETKRTIPAEVRKTLIPEM--P 402
Db 370 YGIGSDDFVCGNSVADADKVKAGQFLNEATADPKQIQIAMPV--NYKSVLGIPWLLIND 428
QY 403 DKQPKQOESL 413
Db 429 NDAKDNKASL 439

RESULT 2
QY 0980L5 PRELIMINARY: PRT: 461 AA.
AC 0980L5:
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE ABC TRANSPORTER XYLOSE-BINDING LIPOPROTEIN.
GN MYPU_3460.
OS Mycoplasma pulmonis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Molllicutes;
OC Mycoplasmatetaceae; Mycoplasma.
OX NCBI_TaxID=2107;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OAB CTIP;
RX MEMLIN=21267165; PubMed=11353084;
R Moszer I., Dybvig K., Wroblewski H., Viart A., Rocha E.P.C.,
Blanchard A.;
RT "The complete genome sequence of the murine respiratory pathogen
Mycoplasma pulmonis."
RL Nucleic Acids Res. 29:2145-2153(2001).
DR EMBL; AL445564; CAC13519.1; -.
DR MyPulast; MYPU_3460; -.
DR InterPro; IPR003760; Bmp.
DR Pfam; PF02608; Bmp; 1.
KW Complete proteome.
SQ SEQUENCE 461 AA; 51096 MW; 6724D4D820809CE4 CRC64;

Query Match 25.6%; Score 552; DB 16; Length 461;
Best Local Similarity 31.2%; Pred. No. 2.9e-24;
Matches 133; Conservative 82; Mismatches 161; Indels 50; Gaps 11;
QY 13 ADONKQJIT--DVSKISGLVNER-----KSEIMAAKADANKHFGINMAIVTAGTV 61
Db 28 AONPKKTSNDSKISITDLSQKVEYETOKIYENKIKQASLETQK-----VTLITADGNI 82
QY 62 NDNSFNGGWEAIIQQLGALTGGEITS-----VDSSTAELEGKYSSILANTNK 107
Db 83 DDKSFNOOVYESOKTLKDFVDKAYKSONKKEANQHKLDNYINSAYKDIIEONKVALDRCY 142
QY 108 NWVYLSGPHODAFTRMLKIPENKOLFTEKNIIILGIDMTDLEN-VIPIGRINILTYKTE 166

```

```

Db 143 TFWILTGFOQGEIENFLNDENNLRFEKNKVIIGVDPANPANSKIFQGSISILFTE 202
QY 167 EAGWLAGYANASFLAKKFPSPDPTKRSATVIGGISPAVTDFIAGYLAGIKAWN--LKNSD 224
Db 203 EAGMGAGVSADPFLGTKVAANNAKRAISAFGGGDFAGVTDFLNGFPEGIRAAVNSAEKAN 262
QY 225 KTKKITTDKIEINLGFVDVODTSTKERLEQIASKDKRSTLLAVAGPLTEIFSDI--ANQ 281
Db 263 KKKIVISENLVYDTGF-IPNAEKNEVSVNVETGKSTISLPVAGDPGTGVVDVLRKQSD 321
QY 282 NDRYLIGVDTQSLVYTKNKFPTSLIKMGYSVFSVLSDLTY-----KKSNSNL 333
Db 322 EDRFVGVDTQSLSTFNDSKRFSTSYKNIAPVYQILLALLTLDESDVILKEGNDKFL 381
QY 334 AGFEFGKSAVYVYLGIGKRFVDIADTSLGENDKRLATEISAISA--KKEFEETKRTIPAE 391
Db 382 GS--NPKNLVYKRLSKAFVNIITSRKVESIKTQADISIQAIKMANPANSKIEREM 438
QY 392 VKRTLE 397
Db 439 TNGDLE 444

RESULT 3
QY 09X775 PRELIMINARY: PRT: 465 AA.
AC 09X775:
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE P48 MEMBRANE LIPOPROTEIN PRECURSOR.
GN P48.
OS Mycoplasma agalactiae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Molllicutes;
OC Mycoplasmatetaceae; Mycoplasma.
OX NCBI_TaxID=2110;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M7;
RX MEDLINE=20002620; PubMed=10531294;
RA Rosati S., Pozzi S., Robino P., Montinaro B., Conti A., Fadda M.,
Pittau M.;
RT "P48 major surface antigen of Mycoplasma agalactiae is homolog to a
malp product of Mycoplasma fermentans and belongs to a selected family
of bacterial lipoproteins."
RL Infect. Immun. 67:6213-6216(1999).
DR EMBL; AJ132423; CAB43718.1; -.
DR InterPro; IPR003760; Bmp.
DR Pfam; PF02608; Bmp; 1.
KW Signal; Lipoprotein.
FT SIGNAL 1 22
FT CHAIN 23 465 P48 MEMBRANE LIPOPROTEIN.
SQ SEQUENCE 465 AA; 51149 MW; 60AD5448CFEB03C56 CRC64;

Query Match 10.1%; Score 391; DB 2; Length 465;
Best Local Similarity 31.2%; Pred. No. 5.9e-15;
Matches 119; Conservative 76; Mismatches 128; Indels 58; Gaps 18;
QY 55 VTAGTIVDNSFNGGWEAIIQQLGALTGGEITSVDSSTAELEGK-----YSSL 102
Db 70 ITDEGSVHDESNOSGWEAVHVSVEIGLDKRAQV--SGKRNLRNRYEPKQQLLEAVYKNA 128
QY 103 ANTNNKVVVLSGFGDAFTRLKIPENK-OLFTEKNIIILGIDWT-----DRENV----- 152
Db 129 IDSGRFYVLCGFTQASL---VGLDENYIKIKDNNTIITVDNLFTEEDANVKRTIK 185
QY 153 -IPTGRYINTLYTTEAGWLAGYANASFLAKKFPSPDPTKRSATVIGGISPAVTDFIAGY 211
Db 186 KIGEGHLVPIVFDTRQAAVYIAGBALADYFSQYXQGPBKRTIGAFGILPWPAVSDFIAGT 245
QY 212 LAGIAMLKLSKDKTKTTTDKIEINLGFVDVODTSTKERLEQIASKDKRSTLLAVAGPL 270

```

DB 246 FGGIDMKNREHPAKTSINTEIENLTLF---TSGTPQATTAINSVVKATASYPVAGSL 302
OY 271 TEIFSDI--IANONDRYLIGVDTDSLYYTKTKN-----KFTSILKNLGYSVSVLSL 323
DB 303 TDPAKIKKILAK-DKFIIGVADQ-----KNALKGRIRFISVKKLIGQANYNIILADL 354
OY 324 YTKKSNSRNUL-AGFEFGKKSAT-VYLGIKD---RFVADIADTS-LEGNDRKILATEAISEA 376
DB 355 YKGEHQDLQPGFEIGKNGKPTVGVGYDTEDKQYGVATSGILDKKDEIANKALKDA 414
OY 377 KKEFEKTKTIPAEVRYKLE 397
DB 415 TAYVYOK-KT---EIQKSLK 430

RESULT 4
052311
ID 052311 PRELIMINARY; PRT; 428 AA.
AC 052311;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
MEMBRANE LIPOPROTEIN P48V.
Mycoplasma fermentans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_Taxid=2115;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PG18;
RA Rawadi G., Dyer K., Dujeancourt A.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF036106; AAB99740.1; -
DR InterPro: IPR003760; Bmp.
DR Pfam: PF02608; Bmp; 1.
KW Lipoprotein.
SO SEQUENCE 428 AA; 47862 MW; D9506E817E30DEAA CRC64;

Query Match 17.8%; Score 384.5; DB 2; Length 428;
Best Local Similarity 27.0%; Pred. No. 1.3e-14;
Matches 116; Conservative 86; Mismatches 157; Indels 71; Gaps 17;
OY 31 NERKSEIMAKADANKH-----FGLMMAIVTAGGVNDNSFNOSGWEAI 74
DB 27 NDNESNISFEKEDISKYTTTNANGQYVKNALKLKLPVLTDEGKIDDSFNOSAPEAL 86
OY 75 QOLGALTGETTSVDSSTAELEKYSLSANTKNVNLVSGFHGDAFTRMKIPENKOLF 134
DB 87 KAIKQKQIEINSVEPS-SNFEASVNSALSAGHKIWLNGFKHQOSIKQY--IDAHEEL 143
OY 135 TEKNIITLIGDMTDENVITPTGRYINLTYKTEEGWLAGANASFLAKKPPSPPTKSAI 194
DB 144 ERNQIKITIGIDF-DIETEKW--FYSLQFNIKESAFITGYAIAWSLSEQ--DESKRYVA 197
OY 195 VIGGGISPAVTDFTAGYLAGIKAMNLKNSDKTKTI-TDKIEINLGFVDVDTSTKERLEQ 253
DB 198 SFGGGAFFPGVTTTNEGFAKILYLNQKH--KSYKFTHTSPVKLDSG-----TAGEMNT 250
OY 254 IAS-----KDKPSTLLAAGPLTEIFSDLIANONDRIYLVDTDSLYYTKTKN 303
DB 251 VINNVLSSTPADYKYNPHVILSVAGPAT--FEVRLANKQGYVIGVDSOGMI--QDKDR 306
OY 304 FFTSILKNLGYSVSVLSL-----YTKKSNSRNLAGFEFGKKSATVYLIGIDRF 353
DB 307 ILTSVLAKHIQAYETLLDLLEKEBEGYKPYVVKDKKADKMSHFQO-----KEKW 358
OY 354 VDIADTSLGNDRKL-ATEAISEAKKEFEKTKTIPAEVRYKLEIPEMPD-KOPDKQOE 411
DB 359 IGVAENHFSMTEDAKINNKEIKAKFKK-----LPEDFYKYINSDKALDKGKIDVSE 414
OY 412 SLDKLITDIN 421
I: : : : : I

DB 415 RLEAIIISAIN 424
RESULT 5
09RGX5
ID 09RGX5 PRELIMINARY; PRT; 428 AA.
AC 09RGX5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE MACROPHAGE ACTIVATING LIPOPROTEIN-404 PRECURSOR.
GN MALP.
OS Mycoplasma fermentans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_Taxid=2115;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SK5;
RX MEDLINE=99115554; PubMed=9916088;
RA Calcutt M.J., Kim M.F., Karpas A.B., Muhlradt P.F., Wise K.S.;
RT "Differential posttranslational processing confers intraspecies
RT variation of a major surface lipoprotein and a macrophage-activating
RT lipopeptide of Mycoplasma fermentans";
RL Infect. Immun. 67:760-771(1999).
DR EMBL; AF099211; AAD16395.1; -
DR InterPro: IPR003760; Bmp.
DR Pfam: PF02608; Bmp; 1.
KW Signal; Lipoprotein.
FT SIGNAL 1
FT CHAIN 25 428 POTENTIAL.
SO SEQUENCE 428 AA; 47835 MW; D03F0F47EA2B1460 CRC64;

Query Match 17.7%; Score 382.5; DB 2; Length 428;
Best Local Similarity 27.0%; Pred. No. 1.6e-14;
Matches 116; Conservative 87; Mismatches 156; Indels 71; Gaps 17;
OY 31 NERKSEIMAKADANKH-----FGLMMAIVTAGGVNDNSFNOSGWEAI 74
DB 27 NDNESNISFEKEDISKYTTTNANGQYVKNALKLKLPVLTDEGKIDDSFNOSAPEAL 86
OY 75 QOLGALTGETTSVDSSTAELEKYSLSANTKNVNLVSGFHGDAFTRMKIPENKOLF 134
DB 87 KAIKQKQIEINSVEPS-SNFEASVNSALSAGHKIWLNGFKHQOSIKQY--IDAHEEL 143
OY 135 TEKNIITLIGDMTDENVITPTGRYINLTYKTEEGWLAGANASFLAKKPPSPPTKSAI 194
DB 144 ERNQIKITIGIDF-DIETEKW--FYSLQFNIKESAFITGYAIAWSLSEQ--DESKRYVA 197
OY 195 VIGGGISPAVTDFTAGYLAGIKAMNLKNSDKTKTI-TDKIEINLGFVDVDTSTKERLEQ 253
DB 198 SFGGGAFFPGVTTTNEGFAKILYLNQKH--KSSKIYHTSPVKLDSG-----TAGEMNT 250
OY 254 IAS-----KDKPSTLLAAGPLTEIFSDLIANONDRIYLVDTDSLYYTKTKN 303
DB 251 VINNVLSSTPADYKYNPHVILSVAGPAT--FEVRLANKQGYVIGVDSOGMI--QDKDR 306
OY 304 FFTSILKNLGYSVSVLSL-----YTKKSNSRNLAGFEFGKKSATVYLIGIDRF 353
DB 307 ILTSVLAKHIQAYETLLDLLEKEBEGYKPYVVKDKKADKMSHFQO-----KEKW 358
OY 354 VDIADTSLGNDRKL-ATEAISEAKKEFEKTKTIPAEVRYKLEIPEMPD-KOPDKQOE 411
DB 359 IGVAENHFSMTEDAKINNKEIKAKFKK-----LPEDFYKYINSDKALDKGKIDVSE 414
OY 412 SLDKLITDIN 421
DB 415 RLEAIIISAIN 424
RESULT 6
032417

ID	Accession	Source	Length	Score	Match	Indels	Gaps
AC	032417	PRELIMINARY	428 AA.				
DT	01-JAN-1998	(TREMBlrel. 05, Created)					
DT	01-JAN-1998	(TREMBlrel. 05, last sequence update)					
DT	01-JUN-2001	(TREMBlrel. 17, last annotation update)					
DE	M161AG.						
OS	Mycoplasma fermentans.						
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;						
OC	Mycoplasmataceae; Mycoplasma.						
OX	NCBI_TaxID=2115;						
RN	[1]						
RP	SEQUENCE FROM N.A.						
RC	TISSUE=BONE MARROW;						
RA	MEDLINE=98022661; Pubmed=9359703;						
RA	Matsumoto M., Takeda J., Inoue N., Hara T., Hatanaka M., Takahashi K.,						
RA	Nagasawa S., Akeho H., Seya T.;						
RT	"A novel protein that participates in nonself discrimination of						
RT	malignant cells by homologous complement.";						
RT	Nat. Med. 3:1266-1270(1997).						
RN	[2]						
RP	SEQUENCE FROM N.A.						
RA	Matsumoto M., Seya T.;						
RL	Submitted (APR-1999) to the EMBL/Genbank/DBJ databases.						
RN	[3]						
RP	SEQUENCE FROM N.A.						
RA	MEDLINE=98241611; Pubmed=9575196;						
RA	Matsumoto M., Nishiguchi M., Kikkawa S., Nishimura H., Nagasawa S.,						
RA	Seya T.;						
RT	"Structural and functional properties of complement-activating protein						
RT	M161AG, a Mycoplasma fermentans gene product that induces cytokine						
RT	production by human monocytes.";						
RL	J. Biol. Chem. 273:12407-12414(1998).						
RN	[4]						
RP	SEQUENCE FROM N.A.						
RA	Seya T., Matsumoto M.;						
RT	"MEMBRANOUS PROTEIN M161AG AND CYCLIC-DNA CAPABLE OF CODING THE						
RT	SAHE.";						
RL	Patent number JP1997157295-A/1,						
RL	17-JUN-1997. KAGAKU GIYUTSU SHINKO JIGYODAN.						
DR	EMBL; D64083; BAA2350.1; -						
DR	EMBL; AB026157; BAA77211.2; -						
DR	InterPro: IPR003760; Bmp.						
DR	Pfam: PF02608; Bmp; 1.						
SO	SEQUENCE 428 AA; 47862 MW; F43B078F21DADD6E CRC64;						

OY	354	VQIADTSLSEGNKKL-ATPAISAEKKEFEERKRTITPAEVRKLTLEIPRPD-KQPDKOGE	411
Dd	359	ICVAAEHESNTSEGOAKINKIKAIKMEKE---LPEDVKYINSDKALKDGKNIDNVSE	414
OY	412	SJDKLTYTDIN 421	
Dd	415	RLEAIIISAIN 424	
RESULT	7		
O9RGX6		PRELIMINARY;	PRT; 429 AA.
ID	09RGX6:		
AC	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)		
DE	MACOPHAGE ACTIVATING LIPOPROTEIN-404 PRECURSOR.		
GN	MALP.		
OS	Mycoplasma incognitus.		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;		
OC	Mycoplasmataceae; Mycoplasma.		
OX	NCB1_TaxId=2101;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-INCOGNITUS.		
RX	MEDLINE=99115554; PubMed=99116088;		
RA	Calcutt M.J., Kim M.F., Kaipras A.B., Muhlradt P.F., Wise K.S.;		
RT	"Differential posttranslational processing confers intraspecies		
RT	variation of a major surface lipoprotein and a macrophage-activating		
RL	lipopeptide of Mycoplasma fermentans.";		
DR	Infect. Immun. 67:760-771(1999).		
DR	EMBL: AF099210: AAD16394.1; --		
DR	InterPro: IPR003760; Bmp.		
DR	Planr: PF02608; Bmp; 1.		
KW	Signal; Lipoprotein.		
FT	SIGNAL 1 24 POTENTIAL.		
FT	CHAIN 25 429 MACROPHAGE ACTIVATING LIPOPROTEIN-404.		
SEQ	SEQUENCE 429 AA; 47961 MW; BD609A812AC3171 CRC64;		
Query Match	17.6%; Score 379; DB 2; Length 429;		
Best Local Similarity	27.0%; Pred. NO. 2.6e-14;		
Matches 117; Conservative	88; Mismatches 151; Indels 78; Gaps		18;
OY	31	NERKSEIMAAKDARKH-----FGLNMAITYAGTVNDNSTSGWEAI 74	
Dd	27	NNDESNISEFKEDIKYTTTNANGROYVKNMALLKLPLLYIDECKIDKSNQSFEAL 86	
OY	75	COLGALUTGEITSVSSPTALEGKYSILANTNKNYWVLSCFOHGAFPMKLIKPENKOLF 134	
Dd	87	KAINOTGTETINNVPSS-SNFESAIVNSALSAGHKTIWLVNGFHHOOSIKOY--IDAHRREL 143	
OY	135	TEKNIIILIGIDMTDENVIPIGRYINLTYKTEAGWLACVANASEFLAKFPSDPTRSAI 194	
Dd	144	ERNQIKIICIDE-DIETEKW--FYSLONIKESAFITGYALAISMLSEQ--DESKRYVA 197	
OY	195	VVGGSIPRAYVDINGIYLAGIKAMWLKNSDKTKI-TDDKIINGSPVOPISTERLEQ 253	
Dd	198	SFGGAFAPEGVTFENSGFAAGILYYNQKH--KSSKIYHTSPVKLDGCF----TADEKMNT 250	
OY	254	IAS-----KKPSYLAVAGPLTEIFEDI-IANONDXYLLGVPTDOSLVYTKTKRN 302	
Dd	251	VANNVLSSTPADYKKNPHVILSYAGPAT--FFTVLAANKGOYVIGVSDOGMT--QDKD 306	
OY	303	KFTSILNLGLSVESVLSL-----YTKSNSRNLGAFEGGKSATVYLGIDKR 352	
Dd	307	RIUTSVLKHIQAAYVEFLDLTLKEDEGGKPPVVADKADKWKMSHPGQ-----KEK 358	
OY	353	FVDIADTSLSEGDKKL-AIEAISSEKKER---EKTITPAEVRKLTLEIPRMDKODP 407	
Dd	359	WIGAIVENHSNTEQAKINKIKAIKMEKELPBEPDVKVYINSDKVLK-----DGNKID 411	

RX MEDLINE=21089007; PubMed=11271425;
 RA Graentzdorffer A., Plich A., Andresen J.R.:
 RT "Molecular analysis of the *grd*-operon encoded proteins of the glycine
 RL reductase and thiorodoxinsystem from *Clostridium sticklandii*,"
 Arch. Microbiol. 175:8-18(2001).
 DR EMBL; AJ276209; CAC14304.1; -;
 DR InterPro; IPR003760; Bmp.
 DR Pfam; PF02608; Bmp; 1.
 DR Lipoprotein.
 SQ SEQUENCE 349 AA; 37083 MW; 53AE4511AFCD5B2E CRC64;

Query Match	10.6%	Score 228;	DB 2;	Length 349;
Best Local Similarity	26.5%	Pred. No. 1e-05;		
Matches 107;	Conservative 65;	Mismatches 138;	Indels 94;	Gaps 21

[illegible]

```

RESULT 13
099ZHA
ID 099ZHA PRELIMINARY; PRT; 350 AA.
AC 099ZHA;
01-JUN-2001 (TREMBLrel. 17, Created)
01-JUN-2001 (TREMBLrel. 17, Last sequence update)
01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PUTATIVE LIPOPROTEIN.
GN SPY1228.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
CC Streptococcus
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SF370 / ATCC 700294 / SEROTYPE M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferretti U.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Seze S., Furvor A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najaf F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes."
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
DR EMBL; AE006563; AAK34087.1; -.
DR InterPro; IPR001114; Adenylsuc_synt.
DR InterPro; IPR003760; Bmp.
DR InterPro; IPR001761; Peripla_BP_like.
DR InterPro; IPR000217; Tubulin.

```

DR Pfam; PF02608; Bmp; 1.
DR Pfam; PF00532; Peripla_BP_like; 1.
DR PROSITE; PS00227; TUBULIN; UNKNOWN_1.
KW Lipoprotein; Complete proteome.
SQ SEQUENCE 350 AA; 36396 MW; 7B5A6A95E0BBFD2 CRC64;

Query Match	10.2%	Score 230.5;	DB 16;	Length 350;
Best Local Similarity	27.8%	Pred. No. 2.8e-05;		
Matches 103; Conservative	49;	Mismatches 150;	Indels 69;	Gaps 17;

[illegible]

RESULT	14			
09CFM9				
ID	09CFM9	PRELIMINARY:	PRT:	350 AA.
AC	09CFM9			
DT	01-JUN-2001	(TREMBLrel. 17, Created)		
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)		
DT	01-OCT-2001	(TREMBLrel. 18, Last annotation update)		
DE	BASIC MEMBRANE PROTEIN A.			
GN	BMPA.			
OS	Lactococcus lactis (subsp. lactis) (Streptococcus lactis).			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;			
OC	Lactococcus.			
OX	NCBI_TaxID=1360;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-IL1403;			
RX	MEDLINE=21235186; PubMed=1137471;			
RA	Boletín A., Wincker P., Mauger S., Jallion O., Malarme K.,			
RA	Weissenbach J., Ehrlich S.D., Sorokin A.;			
RT	"The complete genome sequence of the lactic acid bacterium Lactococcus			
RT	lactis ssp. lactis IL1403."			
RL	Genome Res. 11:731-753(2001).			
DR	EMBL; AE006374; AAK05536.1; -.			
DR	InterPro; IPR003760; Bmp.			
DR	Pfam; PF02608; Bmp; 1.			
KW	Complete proteome.			
SQ	SEQUENCE	350 AA;	36652 MW;	24A30DBEEDD4C814 CRC64;

Query Match	9.4%	Score 203.5	DB 16	Length 350
Best Local Similarity	24.7%	Pred. No. 0.00027		
Matches 97	Conservative 63	Mismatches 157	Indels 75	Gaps 17

```

0Y 20 ITDYSKIGELVNERSETEMAAKKANHFGLNMAIVAGVNDNSPNOGMEILOLA 79
Db 13 LASVAVLACGSHDAAGSGRAKT-----LKAIVTEIGVNDRSNOSAMELOSGWK 66
0Y 80 LTG-----GEITSDVSTAELEGRYSSILANTNKNWVLSGFHOHDATFRLKIPENKOL 134
Db 67 ENNLKKGHYTYFQSNASADYTTNYNSAEGQYKLEFGTFSLQDANSAAK---NN--- 120
0Y 135 TEKNIIILIGIMTENTENVIPGRINLTYTEEGWLAGIYANMSFLAKKFPDPTKRSAL 194
Db 121 PKSNFVIVDSYIKOOKNV-----ASATFADNESAYLAGVAAAK-----ATKTKI 165
0Y 195 -VIGGJSPAYTDFLAGLAGIKAMNLKNSDKKRTITDKIEINLGRVODSTKBELEQ 253
Db 166 GFIGGMOVDYITREKEGYEMAKSVN-----PDIVDOYVAGSFDAKAGKTIAA 215
0Y 254 IASKDRSTLLAVAGPL-TEIFSDIIA-----NONDR-VLIGDVODUSTV--YT-----KT 300
Db 216 AMYGAGDDVVQACGAGVGTGIVFSAPKALNTSKKNADKNVWIGVDODEYIGTKKSKDGKD 275
0Y 301 KNKEFTSLKMLGYSVFSVSLDYTKKSNRNLAG--FEFGKKSATVYLGIKDFVDIAD 358
Db 276 SNFPLVASTIKEVG-----NVWADIKDKTKDGKFPGCTIVTYDLKNGVNLGL-----DSAN 326
0Y 359 TSLGNDKKLATAEISEKKEFEFEKTTITIAE 390
Db 327 SEIK-----DAVAAKADIIDIGKATIPSK 350

```

RESULT	15			
097RHO				
ID	097RHO	PRELIMINARY;	PTT;	350 AA.
AC	097RHO;			
DT	01-OCT-2001 (TEMBLrel. 18, Created)			
DT	01-OCT-2001 (TEMBLrel. 18, Last sequence update)			
DT	01-DEC-2001 (TEMBLrel. 19, Last annotation update)			
DE	LIPOPROTEIN.			
GN	SP0845.			
OS	Streptococcus pneumoniae.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;			
CC	Streptococcus.			
OX	NCBI_TaxID=1313;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=TIGR4;			
RC	MEDLINE=21357209; PubMed=11463916;			
RA	Rettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,			
RA	Peterson S., Heidelberg J., Deboy R.T., Haft D.H., Dodson R.J.,			
RA	Durkin A.S., Glyn M., Kolonay J.F., Nelson W.C., Peterson J.D.,			
RA	Unayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,			
RA	Holteberg E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,			
RA	McDonald L.A., Feldblyum T.V., Angluoli S., Dickenson T., Hickey E.K.,			
RA	Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,			
RA	Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;			
RT	"Complete genome sequence of a virulent isolate of Streptococcus			
RT	pneumoniae".			
RL	Science 293:498-506(2001).			
DR	EMBL: AE007390; AAK74976.1; -.			
DR	TIGR: SP0845.			
DR	InterPro: IPR001114; Adenylysucc_synth.			
DR	InterPro: IPR003760; Bmp.			
DR	InterPro: IPR000217; Tubulin.			
DR	Pfam: PF02608; Bmp. 1.			
DR	PROSITE: PS000227; TUBULIN; UNKNOWN_1.			
CM	Complete proteome.			
QO	SEQUENCE 350 AA; 36746 MW; CBD6767B7FE501E69 CRC64;			

Query Match	9.48;	Score 202;	DB 16;	Length 350;
Best Local Similarity	25.78;	Pred. No. 0.00033;		
Matches 98;	Conservative 55;	Mismatches 145;	Indels 84;	Gaps 19;

QY 33 RKSEIMAKADANKHEGLNMAIVTAGTVNDNSFNQSGWEAIQQLGA--LTCGETSVD 89

```

Db      25 RSRNRMASSSDYK-----TKAAIYDIDGCVDDKSTNSQAMBELOAMGEHNLISKDNGFTYF 80
QY      90 SSTAE-----GKYSILANTNNKWWLGSFOHGDAFTTBLKJPEKNOLETTEKNI 140
Db      81 QSTSEADVANNLOQAAGSYNLIIEGV-----GFALNMA-----VKDAKEHTDILNTY 126
QY      141 ILIGIDWTDTEENVILPTGRYINLTYTEEBAGWLAGYANASFLAKRPPSDPTKSAIYIGGI 200
Db      127 LIDDIYKQKQNV-----ASVTFADNDSGYLAGVAAK-----TTTKQGVFPYGI 171
QY      201 -SPAVTDFIACYLAGIKAMNLKNSDDKTKTITTDKIEINLGPVDVODYSTKERLEQIASDK 259
Db      172 ESEVISREAFAGKAGV-----ASVDPSIKVOVDYAG-SFG-DAAKGRTIAAQAAGAD- 223
QY      260 PSTLLAANGP-TEIFSDIINQDR-----YLGVDTQSL-VYT-----KTKNKEFT 306
Db      224 --IYQVAVAGGTGACVFEAKSLINESRENEKVMYIVGIDRQDEAGKTKSKDGESNPLV 281
QY      307 SILNLGVSFVSJLDLYTKSKNSRNLAGFEFGKSKATVYLGIKDRVVDIADTSLBENDK 366
Db      282 STLKOVGTIYVDI-----SNKAERGEPPGGQ--VIYYSLKDKGVDAVYTNLSEEGK 330
QY      367 KLAETAISEAKKEFEKTKTIP 388
Db      331 K----AVEDAKAKILLDSGVKVP 348

```

Search completed: July 18, 2002, 15:31:18
Job time: 125 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2002, 09:14:54 ; Search time 22.25 Seconds

(without alignments)
464,361 Million cell updates/sec

Title: US-09-676-249A-4

Perfect score: 1 MMDKTTKEKESADNKNQI.....KDPKQDESLDKLITDINNL 423

Sequence:

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Minimum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
4: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCITUS_COMB.pep:*
6: /cgn2_6/prodata/2/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	202	9.4	328	4	US-08-961-083-8
2	178	8.3	341	1	US-08-396-957A-5
3	161	7.5	339	1	US-08-396-957A-4
4	143	6.6	353	4	US-08-313-412-1
5	137.5	6.4	889	4	US-09-336-447A-15
6	133.5	6.2	1786	4	US-08-973-462-8
7	119	5.5	2285	4	US-09-308-375-2
8	117.5	5.5	1098	1	US-08-409-995-2
9	117.5	5.5	1098	3	US-08-685-467-2
10	117.5	5.5	1098	4	US-09-377-155-32
11	117.5	5.5	1098	4	US-08-913-942-2
12	117.5	5.5	1098	4	US-09-669-974-32
13	117.5	5.5	1098	4	US-08-268-347-44
14	116.5	5.4	1104	4	US-08-923-992A-4
15	113.5	5.3	941	4	US-09-074-658-75
16	113	5.2	2353	4	US-09-377-155-33
17	113	5.2	2353	4	US-08-913-942-4
18	113	5.2	2353	4	US-09-669-974-33
19	112	5.2	2354	4	US-09-268-347-47
20	112	5.2	2411	4	US-09-268-347-36
21	110.5	5.1	1164	4	US-08-923-992A-2
22	109	5.1	3418	2	US-08-603-753D-4
23	109	5.1	3418	3	US-08-755-587-44
24	109	5.1	3418	4	US-09-099-753-4
25	109	5.1	3418	4	US-08-986-106-4
26	108.5	5.0	686	4	US-09-368-169-8
27	108.5	5.0	873	4	US-09-336-447A-13

28	108	5.0	2329	3	US-08-755-587-16	Sequence 16, Appl
29	108	5.0	3418	2	US-08-639-501-2	Sequence 2, Appl1
30	108	5.0	3418	3	US-09-044-946-2	Sequence 2, Appl1
31	108	5.0	3418	3	US-09-044-908-2	Sequence 2, Appl1
32	107.5	5.0	1964	2	US-08-790-912-3	Sequence 3, Appl1
33	107.5	5.0	2052	2	US-08-790-912-2	Sequence 2, Appl1
34	107	5.0	607	1	US-08-409-995-6	Sequence 6, Appl1
35	107	5.0	607	3	US-08-685-467-6	Sequence 6, Appl1
36	107	5.0	607	4	US-08-913-942-6	Sequence 6, Appl1
37	107	5.0	703	3	US-08-646-273-25	Sequence 25, Appl
38	107	5.0	753	3	US-08-646-273-36	Sequence 36, Appl
39	107	5.0	1579	3	US-08-755-587-184	Sequence 184, App
40	107	5.0	1912	1	US-08-409-995-4	Sequence 4, Appl1
41	107	5.0	1912	3	US-08-685-467-4	Sequence 4, Appl1
42	106.5	4.9	892	4	US-09-336-447A-5	Sequence 5, Appl1
43	106.5	4.9	1098	2	US-08-923-992A-8	Sequence 8, Appl1
44	106.5	4.9	1529	4	US-08-728-470-10	Sequence 10, Appl
45	106.5	4.9	1529	4	US-08-719-641-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-08-961-083-8
; Sequence 8, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; City: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 Inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 328 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-961-083-8

Query Match 9.4%; Score 202; DB 4; Length 328;
Best Local Similarity 25.7%; Pred. No. 5.6e-11;
Matches 98; Conservative 55; Mismatches 145; Indels 84; Gaps 19;
OY 33 RKSTMAKADKANKHGLNMAITYAGCTVDNNSNGWEAIDQLGA---LTGGETISVD 89

MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/396,957A
FILING DATE: 01-MAR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/020,245
FILING DATE: 19-FEB-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/664,731
FILING DATE: 05-MAY-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/487,716
FILING DATE: 05-MAY-1990
ATTORNEY/AGENT INFORMATION:
NAME: DOROTHY R. AUGH
REGISTRATION NUMBER: 36,434
REFERENCE/DOCKET NUMBER: 2026-4018US4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEO ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 339
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Borrelia burgdorferi
STRAIN: Sh-2-82
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
FEATURE:
NAME/KEY: p39
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION: p39, protein
OTHER INFORMATION: sequence.
08-396-957A-4

Query Match 7.5%; Score 161; DB 1; Length 339;
Best Local Similarity 22.8%; Pred. No. 4.5e-07;
Matches 85; Conservative 68; Mismatches 136; Indels 84; Gaps 18;

09 GTVNDSPNOSGWEALIOGLALGELTSDST---ALEKRYSLANTNNWVLG 114
DB 37 GTFDDSPNLSALNGVAKAEKIEVLAKESSNSTYLSLEGG---LKNAGSDLLWLG 92
0Y 115 FQHG-D-AFTRWLKPENKOLFTEKNIILIGIDMTDENVIPTGRVYINLTVKTEAGWLAG 173
DB 93 YRSVDVAKVAALQNPMPKVAI-----IDPIYSNDPIR-ANLVGMFRAEGAFLENG 142
0Y 174 YANASFLAKFPSPDKRSAYIVGGISPAVVD-FTAGYIAGIKANWLKSKDKTKITTD 232
DB 143 YIAAKL-----SKTGKIGLGIIEGEIVDAFRGYEGAKAYAN---KDIXISTQ 188
0Y 233 KI-----EINLGFVDVDTSTKERLEQIASKDKPSTLLAVAGPLTEIFSDITANO--NDRYL 286
DB 189 YISFADLEAGRSVATRMYSDEID-----IHHAGLGIGCAIEVAKELGSGHYI 238

0Y 287 IGVDTDOSLVYTKTKNFKFTSLIKNLGYSVFVSLDLYTKRSNRNLAGEFEGKKSATVY 346
DB 239 IGVDEDOAVL---APDVITSTTKDVGRL-----NIFT--SHNLKTNFEGSK---LIN 285
0Y 347 LGIKDRFVDIADTSLEGNKKLATFALISEAKKEFEETKTIPAEVKTILEPMDKOP 406
DB 286 YGLKEGVGVFV-----RNPWMISF-----ELEKIDMLSKKIINKELIIV----- 325
0Y 407 DKQOESLDKLTID 419
DB 326 -SMKESEKFLKE 337

RESULT 4
US-08-313-412-1
Sequence 1, Application US/08313412
Patent No. 6248583
GENERAL INFORMATION:
APPLICANT: Aron Ph.D., Lieselotte
APPLICANT: Cabello M.D., Felipe
APPLICANT: Godfrey M.D., Henry P.
APPLICANT: Schwartz Ph.D., Ira
TITLE OF INVENTION: Chromosomally-Encoded Membrane
NUMBER OF SEQUENCES: 2
TITLE OF SEQUENCES: Protein of Borrelia burgdorferi
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle
STREET: P.O. Box 1051, Clinton Square
CITY: Rochester
STATE: New York
COUNTRY: USA
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/313,412
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 35553/1020
TELECOMMUNICATION INFORMATION:
TELEPHONE: 716-263-1304
TELEFAX: 716-263-1600
INFORMATION FOR SEO ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 353 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-313-412-1

Query Match 6.6%; Score 143; DB 4; Length 353;
Best Local Similarity 23.4%; Pred. No. 2.4e-05;
Matches 83; Conservative 54; Mismatches 131; Indels 86; Gaps 18;

0Y 55 VTAGGTVDNDSFNOSGWEALIOGL-----ALTGELTSDVSTAELEGKY 99
DB 34 VLAHGSFYDYGIVDSYDVGVKLRDNFGILITKSLRPYIEGRRLTVDE--AMTEDAY 91
0Y 100 SSLANTNNWVLSGFGHGAFTRWLKPENKOLFTEKNIILIGI-----DWTDTENVIPT 155
DB 92 EVQKNPLMLFMLI-GYRFSDL-----SVKLSYERPDYIYGIIADAFYGDIO--VPR 139
0Y 156 GRINLNTYKTEEGWGLYANASFLAKFPSPDKRSAYIYG---GGISPAVVDFTAGYL 212
DB 140 NS-LAIKFRNEEAFLAGYIAAKMSRKE-----KIGFLTGPMSHLDKFRGFK 187

```

OY      213 AAIKAMN--LKSNDKKTKITTTDKIEINLGFDVQDSTKRLEDA----SKMPSFLAV 266
Db      168 AGIFANPRRLRYLSKKAPSLF-----KKGKMALEFKREDKVYGFIP 232
OY      267 ACPLTEIFSDIANONDR-YLIGVDTDOSLVYTKRKNEFTSILKNAGYSVSFLSDLYT 325
Db      233 AGITGICGVADAKELCPKYVIGLNDQSYI---APQNVITSIIKDGVYSISSEYI- 288
OY      326 KKSNSRNLAGFEFGKKSATVYLGINDFPDIA-DTSLEGNDKKLATEALSEAKR 378
Db      289 ---NNRVKFG-----GIIDRGLEGVETIVKDPDVANN--RLVDEVIDLENK 331

RESULT          5
US-09-336-447A-15
; Sequence 15, Application US/09336447A
Patent No. 6310190
GENERAL INFORMATION:
APPLICANT: HANSEN, ERIC J.
APPLICANT: AEBI, CHRISTOPH
APPLICANT: COPE, LESLIE D.
APPLICANT: MACIVER, ISOBEL
APPLICANT: FISKE, MICHAEL J.
APPLICANT: FREDENBURG, ROSS A.
TITLE OF INVENTION: USP41 AND USP42 ANTIGENS OF MORAXELLA CATARRHALIS
FILE REFERENCE: AMCY:024
CURRENT APPLICATION NUMBER: US/09/336,447A
NUMBER OF FILING DATE: 1999-06-21
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 15
LENGTH: 889
TYPE: PRY
ORGANISM: Moraxella catarrhalis
US-09-336-447A-15

Query Match           6.4%; Score 137.5; DB 4; Length 889;
Best Local Similarity 22.8%; Pred. No. 0.00036;
Matches 112; Conservative 78; Mismatches 198; Indels 103; Gaps

OY      16 ONKOITDVSKISGLVN---ERKSEIMAAKADANKHFGLMALVITAGTVNDSFNOS--- 69
Db      210 ONEASGDSTRSVSGGYNNLAEGKSSAIG-GEFNALGNMATI--GGGRONEASGDRSTVA 266
●       70 GWEALIOQG---ALNG--EITSVDSRA-----ELEBKYSILA----- 103
Db      267 GGEONOLAIKYSTTGSCQRONEASGDRSTVAGEONOLAIKYSTVSGCYRNAOTGKGSFAA 326
OY      104 ----WTNKNVWLSEFOHGDAFTRMUKTIPENROL-FTEKNITIIIGIDMTDENVIPTGRY 158
Db      327 GIDNKANMDNMVALGNKNTIEGENSVALIGSNNTYVKKNQKNVFITLSN-IDTIDA-QSGSV 384
OY      159 I---NLTYKTEBA-----CWLAYANASFLAKKPSPDPKRSIAIVTGG-ISPA 203
Db      385 LIGDMTSCRKAATAVEDATVGDLSLTGFAGVSKANSGTIVSVEGGERQIVHYGAGRISND 444
OY      204 VTDFLAGVLAKKANLNKNSDDKKTITTDKIEINLGFDVQDSTKERLEQIASXDKPSTL 263
Db      445 STDANNGSQLVALAAYVDNOYDIERNODDLAKNQADIAKNADIAKNADIoTLENDVGKE---L 500
OY      264 LAVAGPLETFEISDIANONDRYLIGVDTDOSLVYTKRKNEFTSILKNAGYSVSFLSDL 323
Db      501 LNLSGRLIDOKRADIDONNIINHIELAQODOSHSDIKT-----LKNVEBELLLESHL 553
OY      324 YTKSNS-SRNLAGFEFGKKSATVYLG--IKDRFYADIADTSLSEGNCKLA----- 369
Db      554 IDOKADTLKRDIALAESNVEEGLLDSLGRLIIDQKADIAQN--CANIIDLAAYNELDOYAO 611
OY      370 --TEAISAKKEFEKTKTIPEAEVRKTLLEIPEMDDKQPKQOESLDK----- 415
Db      612 KOTEIIDLAKNASSENTONTI--EDIAAYNELADAVAKOOTEIIDLAKNASSENTONTIANKN 669

```

```

OY      416  ---LITDINNL 423
          :  :|||:
Db      670  QADIANNINNI 680

RESULT  6
US-08-973-462-8
: Sequence 8, Application US/08973462B
: Patent No. 6191270
: GENERAL INFORMATION:
: APPLICANT: DRULIHE, PIERRE
: APPLICANT: DAUBERSIES, PIERRE
: TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
: FILE REFERENCE: 0660-0125-0 PCT
: CURRENT APPLICATION NUMBER: US/08/973,462B
: CURRENT FILING DATE: 1998-02-06
: EARLIER APPLICATION NUMBER: PCT/FR96/00894
: EARLIER FILING DATE: 1996-06-12
: EARLIER APPLICATION NUMBER: FR 95/07007
: EARLIER FILING DATE: 1995-06-13
: NUMBER OF SEQ ID NOS: 29
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 8
: LENGTH: 1786
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: Polypeptide
US-08-973-462-8

```

```

Query Match 5.2%; Score 133.5; DB 4; Length 1786;
Best Local Similarity 19.3%; Pred. No.0.0027;
Matches 85; Conservative 77; Mismatches 145; Indels 133; Gaps 19;

QY 9 EKSADNONKOITDYSKISGLVNERKSEI-----MAAKADANKRGLMAIVTAGG 59
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 908 ENTQSEEEKKEIVDY-----IEEVKEAVATLTIEVQAEKRSAN-----TITEIFE 954
QY 60 TVNDNSF--NOSGMAIQOGLALGTGETTSDSSIAIEEGKYSLSLANRKNKVVLSGRFH 117
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 955 NLEAEVASENENVAENLEKRLNEVFNVLDRKVEETVEISGE--SLENNEM----- 1002
QY 118 GDAFTWRLKIPENKOLFERNKNIILIGIDMTDENTENIPGRY---INTKTEAGWLAG 173
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1003 -----DKAFEEIFDNKVGIG-----ENLL-TGMRSITSITYOSE----- 1039
QY 174 YANASFLAKKPPSDPTKRSIAIVIGGIGISPAVTDIAGYIACI--KAMNLIKNSDKKTKITTT 231
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1040 -----KVDLENVYSSILDNIENKKEGLLNKLKNLKNISSTEGVQEVYT 1080
QY 232 DKIEINLGFVDQDPTSTKRELEQIASKDKPSTLLAAGVLTLEF---SPITANQNDRYLIG 288
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1081 EHVEDQNVVDY-----DVPAMKDDPGLIGLNEAGLKGKMFENLDEVPKSESQ--VIT 1129
QY 289 VDTPOSLVYTKTKNKEFTSILKNLGYSVFSYSLDLYTKKSSNRNLAGEPFGRKASTYVLG 348
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1130 VEEIKDEYQKEVEKEYSIIEEMENIVDYLEE---EKED----- 1167
QY 349 IKDRVEDIADTSLF-GNDKRLATEAISBAKKE---PEEKTXTIIPAEVARKTLIEIPMPD 403
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1168 LTDKMIDAVESIEISDSKREETESIKRKEDVSLVAEEVQNDMDDESEKVLKKN-- 1225
QY 404 KQPKQKQESLKLITDINNL 423
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1226 -----EEELMKDAVEINDI 1239

RESULT 7
US-09-308-375-2
Sequence 2, Application US/09308375
Patent No. 6300117

```

```

; GENERAL INFORMATION:
; APPLICANT: Genencor International, Inc.
; TITLE OF INVENTION: Proteases from Gram-Positive Organisms
; FILE REFERENCE: GC94-PCT
; CURRENT APPLICATION NUMBER: US/09/308,375
; CURRENT FILING DATE: 1999-05-14
; EARLIER APPLICATION NUMBER: EP9719636.4
; EARLIER FILING DATE: 1997-09-15
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 2285
; TYPE: PRT
; ORGANISM: Bacillus subtilis
; US-09-308-375-2

```

```

Query Match          5.5%; Score 119; DB 4; Length 2285;
Best Local Similarity 18.9%; Pred. No. 0.096;
Matches 99; Conservative 82; Mismatches 184; Indels 158; Gaps 22;

```

```

10 EKSAONKQITDVSKISGIVNERKSEIMAAKADANKHFGNLMAIYTA---GGTVADNSF 66
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
466 DKINEVDNNYAVTTLDLANSIRKAGSTASTFGEVLENDLIGYTAIASTRESGNIVGNSL 525
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
67 -----NSQGEAIOOLG---ALTGEITVSSTAELEGKYSILANTNK----- 107
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
526 KITPARIKNNQSSIKALEQIGISVKTAGGEAKSASDLISEVAGKWPILSDAKQNTSISG 585
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
108 -NWWVLGTF-----OHGDAFTFMKIPENKOL-----FTEKNT 139
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
586 AGIYQSRFNAMNMFNSIAONAKTAASTGSAWSEQKYADSLQARVNKLQNNFTE--F 643
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
140 IILGIDMTDENVI---PGRYINL-TYKTEBAGW---LAGYANASFLAKFPSPDPK 190
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
644 AIAASAFITSDGLEFTQAAGSLNASTGVKSGFLPLLAIVSTATILLST--WTRL 701
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
191 RSAIVIG-----GGISPA-----VTDFIAGYLAGIKAMNL 220
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
702 ASLLIICTRAMQETIATAGLEAGMTRAAVASHVLTALRGILVSTLVGCAFAAL--GMAL 760
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
221 K--NSDKKTKITTDKIEINLGFVDVDTSTKERLEQIASKDKPSTLLAAGPLTEIFSDI 277
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
761 ESTLISFAEAKKAKDDFEQSOQTNEAITYTNK-----DSTDK-----LIQYKEL 805
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
278 IANQNRVYLGVDFDSLVYTKTKNKFPSILKNLGSVFSYSDLTYSKSNRNLAGFE 337
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
806 QVKESRSILTSDEQETLYDTQQLAQTFFPALVGYDSQGNAIL-----KTNKELEKATE 859
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
338 FGKRSATVYIGIDRFVADIADTSLSEGNDKLATEAISEAKKEFEKTKTI--PAEEVRYKT 395
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
860 NTKS---YIALKKQ-----ETRDSAKKTFEDASKELKKSKDELKQY 897
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
396 LEIPEMPDK-----OPDKQGESLDKLITDINN 422
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
898 KQIADYNDRKGRPKMDLIADDDYKVAADKAKQGLKAKQSDIES 940
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT      8
US-08-409-995-2
; Sequence 2, Application US/08409995
; Patent No. 5646259
; GENERAL INFORMATION:
; APPLICANT: Barenkamp, Stephen I.
; TITLE OF INVENTION: Haemophilus Adhesion Proteins
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albrighton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA

```

```

; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/409,995
; FILING DATE: 24-MAR-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-61053/REF
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1098 amino acids
; TYPE: amino acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; US-08-409-995-2

```

```

Query Match          5.5%; Score 117.5; DB 1; Length 1098;
Best Local Similarity 19.1%; Pred. No. 0.04;
Matches 112; Conservative 75; Mismatches 170; Indels 229; Gaps 28;

```

```

7 TREKESADONKQITDVSKISG-----LVNERKSEIMAAKA--DANKHGLN- 51
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
289 TVEFLSADPTTIVYDSEKENGKRTVEKIGAKTSVKEKNGKLFTEKANKETKAVGCANA 348
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
52 -----MAIVTAGVYVNDNSFNQSGMEAIQOLGALTGGETVSST---AELEGKYS 100
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
349 TEDADEKGLVTFADYI--DAVNKTGWRIKTTDANGQNGDFATVAGSTNVTFAAGNCTTA 406
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
101 SLANTKKNWVVLGSPQHD-----AFTRMVKI-----PENKQ 132
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
407 TVTNGVDGIVKYDAKQGLDKLIDGKIADDTVALTVNDGKANNPKQYADVASTDERR 466
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
133 LFEKNTII--ILGIDMTDT-----ENVLPGRYI-----NLTYKTEBAGW 170
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
467 LVYAKGLVYALNLSWTTTAAEADGTLDSNABEQEVKADKYTEKAGKMLKVKQGBANF 526
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
171 LAGYANASFLAKFPSPDPKRSALVIGGIGSPAVTDF----- 207
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
527 TVSLQDAL-----TGLTSTITLGTGNNGAKTEINKDGLTTPANGAGANNANTISV 576
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
208 -----IAGYLAGIKAW-----NL--KNSDKKT 227
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
577 TKDGISAGGSVKNVYSGLKKFEDANFDPLTSSADNLTKNDDAYKGLNLDKRGTDKQY 636
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
228 KITTDKIEINLGFVDVDTSTKERLEQIASKDKPSTLLAAGPLTEIFSDIINQND--RYL 286
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
637 PVVADNTAAIVG--DLRG-----LGWVISADK-----TTGGSSTG--YHDQVRANNEKFK 662
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
287 IG-----VDTGDSLVYTKTK-----NKEF-----TSILKNLGSVFSYSDLT 333
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
683 SGNGINVGKTVNGRREITFELAKGEVYKSNFEVTKETNGKETSIVK-----VGDK 733
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
324 YTKKSNRNLAG---PEFGKRSATVYIGIDRFVADIADTS-----LEGN----- 364
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
734 YYSKEDIDLTGGPKLKGDTVAARKYODKGGKVAVSVDNTEATITTKGSGSYVIGNOVADA 793
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
365 -----DKLATEAISEAKKEFEKTKTIPEAEVRYKLEIPEMPDK 404
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
794 IAKSGFELGLADEA--DAKRAFDDKTKALSA-----GTTTEIVNADHK 833
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 9

US-08-685-467-2
; Sequence 2, Application US/08685467
; Patent No. 6060059
; GENERAL INFORMATION:
; APPLICANT: St. Gene III, Joseph W.
; APPLICANT: Barenkamp, Stephen J.
; TITLE OF INVENTION: HAEMOPHILUS ADHESION PROTEINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: Flehr, Holbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/685,467
; FILING DATE: 22-JUL-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/409,995
; FILING DATE: 24-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-61053-2/RPT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEEX: 910 277299
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1098 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
US-08-685-467-2

Query Match 5.5%; Score 117.5; DB 3; Length 1098;
Best Local Similarity 19.1%; Pred. No. 0.04;
Matches 112; Conservative 75; Mismatches 170; Indels 229; Gaps 28;
Db 7 TKEESADNOKQITDVSKISG-----LVNERKSEIMAAKA--DANKHFGIN- 51
Db 289 TVEFLSADTETTTVVDSEKNGKRTVEKIGAKTSVIEKEDKLTFGANKKETNKVDGANA 348
QY 52 -----MAITAGGVNDNSFNOSGWEAIOALGALTGETISVDSST---AELEGKYS 100
Db 349 TEDADEGGLVTAKDVI--DAVNKTGWRIKTTDANGOGDPAIVASGNTVFASGNGTTA 406
QY 101 SLANTNKNVWVLSGFQHD-----AFTRWLKI-----PENKQ 132
Db 407 TVTNGTDGITVYDAKVDGDLKLDGKIAADTTALTVDGNKANNPKGVADVASTDEKK 466
QY 133 LFTENKIT--ILGIDWMT-----ENVITPTGRI-----NLTYTEEGW 170
Db 467 LVYAKGLVTALNSLWTTTAAEADGGLDGNASDEYKAGDKVYFKAGKNLKVQEGANF 526
QY 171 LAGYANASFLAKFPSPDPTKRSATVIGGISPAYTDF----- 207
Db 527 TYSLODL-----TGLTSTLGTGNGGAKTEINKDGLTTPANGAGANNANTISV 576
QY 208 -----IAGYLAGIKAM-----NL--KNSDKKT 227
Db 577 TKDGISAGGQSVKNVVSGLKKFKGDANFDP LITSADNLTKONDAYKGLTINDEKGTDKQT 636

QY 228 KITTDKIEINIGFDVQDSTKERLEQIASKDKPSTLLAVAGPTEIFSDIJANQND-RYL 286
Db 637 PVVADNTAATVG-DLRG-----LGMVISADK-----TTGGSTG-YHDQVRANAEVFK 682
QY 287 IG-----VDTOOSLYVYTK-----KKEF-----TSLKRLGVSFVSVDL 323
Db 683 SGNINGVSGKTVNGRREITFELAKGEVYKSNBEYKETNGKETSIVK-----VQDK 733
QY 324 YTKKSNRRLAG---FERGKSAVYVLGIKDRFVADIPTS-----LEGN----- 364
Db 734 YSKREDIDLTTGQPKLKGNVTYAAKYQDKGKVSVTNTENTITNKSGVYTGQVADA 793
QY 365 -----DKKLTAEISEAKKEFEKTKTTPAEVRLTLEIPMDPK 404
Db 794 IAKSGFELGLADEA--DAKRAFDDKTKALSA-----GTTEIVAHDK 833

RESULT 10
US-09-377-155-32
; Sequence 32, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
; APPLICANT: PEAK, Ian Richard Anselm
; APPLICANT: JENNINGS, Michael Paul
; TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
; FILE REFERENCE: 065064/0128
; CURRENT APPLICATION NUMBER: US/09/377,155
; PRIOR FILING DATE: 1998-08-19
; PRIOR APPLICATION NUMBER: PCT/AU98/01031
; PRIOR FILING DATE: 1998-12-14
; PRIOR APPLICATION NUMBER: GB 9726398.2
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 32
; LENGTH: 1098
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-377-155-32

Query Match 5.5%; Score 117.5; DB 4; Length 1098;
Best Local Similarity 19.1%; Pred. No. 0.04;
Matches 112; Conservative 75; Mismatches 170; Indels 229; Gaps 28;
Db 7 TKEESADNOKQITDVSKISG-----LVNERKSEIMAAKA--DANKHFGIN- 51
Db 289 TVEFLSADTETTTVVDSEKNGKRTVEKIGAKTSVIEKEDKLTFGANKKETNKVDGANA 348
QY 52 -----MAITAGGVNDNSFNOSGWEAIOALGALTGETISVDSST---AELEGKYS 100
Db 349 TEDADEGGLVTAKDVI--DAVNKTGWRIKTTDANGOGDPAIVASGNTVFASGNGTTA 406
QY 101 SLANTNKNVWVLSGFQHD-----AFTRWLKI-----PENKQ 132
Db 407 TVTNGTDGITVYDAKVDGDLKLDGKIAADTTALTVDGNKANNPKGVADVASTDEKK 466
QY 133 LFTENKIT--ILGIDWMT-----ENVITPTGRI-----NLTYTEEGW 170
Db 467 LVYAKGLVTALNSLWTTTAAEADGGLDGNASDEYKAGDKVYFKAGKNLKVQEGANF 526
QY 171 LAGYANASFLAKFPSPDPTKRSATVIGGISPAYTDF----- 207
Db 527 TYSLODL-----TGLTSTLGTGNGGAKTEINKDGLTTPANGAGANNANTISV 576
QY 208 -----IAGYLAGIKAM-----NL--KNSDKKT 227
Db 577 TKDGISAGGQSVKNVVSGLKKFKGDANFDP LITSADNLTKONDAYKGLTINDEKGTDKQT 636
QY 228 KITTDKIEINIGFDVQDSTKERLEQIASKDKPSTLLAVAGPTEIFSDIJANQND-RYL 286
Db 637 PVVADNTAATVG-DLRG-----LGMVISADK-----TTGGSTG-YHDQVRANAEVFK 682

OY 287 IG-----VDTOSLVYTKT-----NKEF-----TSILNLSYVFSVLSL 323
 DB 683 SGNGINVSCTVNGRREITFELAKGEVKSNETVAKETNGKETSIVK-----VGDK 733
 OY 324 YTKKSNRLAG---PEFGKKSATVYLGIDRFVDIADTS-----LEGN----- 364
 DB 734 YSKEDIDLTGQPKLKDNTVAKYODKGGKVSVDNTTEATITNKGSGYVGNVADA 793
 OY 365 -----DKLATEPAISEAKKEFEKTKTIPAEVKTLEIPEMDK 404
 DB 794 IAKSGFELGLADEA--DAKRAFDDKTKALSA-----GTEIVNADHK 833

RESULT 11
 US-08-913-942-2
 ; Sequence 2, Application US/08913942
 ; Patent No. 6200578

GENERAL INFORMATION:
 APPLICANT: St. Geme, Joseph
 APPLICANT: Barenkamp, Stephen J.
 TITLE OF INVENTION: HAEMOPHILUS ADHESION PROTEINS
 NUMBER OF SEQUENCES: 19
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Flehr Hobdach Test Albritton & Herbert LLP
 STREET: Four Embarcadero Center, Suite 3400
 CITY: San Francisco
 STATE: California
 COUNTRY: United States
 ZIP: 94111-4187

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/913,942
 FILING DATE: 29-DEC-1997
 CLASSIFICATION: 514

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/409,995
 FILING DATE: 24-MAR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US96/4031
 FILING DATE: 22-MAR-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Vance, Dolly A.
 REGISTRATION NUMBER: 39,054
 REFERENCE/DOCKET NUMBER: A-61053-1/RFT/RMS/DAY

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 781-1989
 TELEFAX: (415) 398-3249
 TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1098 amino acids
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 US-08-913-942-2

Query Match 5.5%; Score 117.5; DB 4; Length 1098;
 Best Local Similarity 19.1%; Pred. No. 0.04; Indels 229; Gaps 28;
 Matches 112; Conservative 75; Mismatches 170;

OY 7 TREKESADNONKOITDVSKISG-----LVNERKSEIIMAAKA--DANKHFGIN- 51
 DB 289 TVEFLSADTETITVYDSKENGKRETEVIGAKTSVIEKDGKLTGTGANKNETNVDGANA 348
 OY 52 -----MAIVTAGYVNDNSFNOSGWEAIOQALGTGETTSVDSST-----ALEBKYS 100
 DB 349 TEDADEBKGIVTAKDVI--DAVNKTGMRIKTTDANGONGDFATVAGSTNVTAFASGNGTTA 406

DB 349 TEDADEBKGIVTAKDVI--DAVNKTGMRIKTTDANGONGDFATVAGSTNVTAFASGNGTTA 406
 OY 101 SLANTNKNVWVLSGFOHGD-----AFTBMLKI-----PENKO 132
 DB 407 TVTNGTDGIVTKDAKAYDGDLDKIDKRIADITLALYVNDKNNANPKGKADVASIDEKK 466
 OY 133 LFTENKII--ILGIDWTD-----ENVIPGRYI-----NLTKTEBAGW 170
 DB 467 LVTAAGLVTLNLSMTTAAEDGGTLDGNASEOEVKADKVTFRAGKNLKVKEGANGF 526
 OY 171 LAGYANASFLAKKFPSPDPTKRSALVIGGIGSPAVTDF----- 207
 DB 527 TVSLADAL-----TGLTSITLGTGNGAKTEINKDGLTTPANGAGANNANTISV 576
 OY 208 -----IAGYLAGIKAW-----NL--KNSDKKT 227
 DB 577 TKDGISAGOSVKNVYSGLKFFGDANFDPLTSSANDITKONDAYKGLTULDEKGDJKT 636
 OY 228 KITTDKIEINLGFVDVDTSTKERLEQJASKDKPSTLLAVAGPLTEIFSDIIANQND-RYL 286
 DB 637 PVVADNTAATVG-DLRG-----LGWITSADK-----TTGGSSTF-YHDQVRNANEVYKFK 682
 OY 287 IG-----VDTOSLVYTKT-----NKEF-----TSILNLSYVFSVLSL 323
 DB 683 SGNGINVSCTVNGRREITFELAKGEVKSNETVAKETNGKETSIVK-----VGDK 733
 OY 324 YTKKSNRLAG---PEFGKKSATVYLGIDRFVDIADTS-----LEGN----- 364
 DB 734 YSKEDIDLTGQPKLKDNTVAKYODKGGKVSVDNTTEATITNKGSGYVGNVADA 793
 OY 365 -----DKLATEPAISEAKKEFEKTKTIPAEVKTLEIPEMDK 404
 DB 794 IAKSGFELGLADEA--DAKRAFDDKTKALSA-----GTEIVNADHK 833

RESULT 12
 US-09-669-974-32
 ; Sequence 32, Application US/09669974
 ; Patent No. 6331173

GENERAL INFORMATION:
 APPLICANT: PEAK, Ian Richard Anselm
 APPLICANT: JENNINGS, Michael Paul
 APPLICANT: MOXON, E. Richard
 TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
 FILE REFERENCE: 065064/0128
 CURRENT APPLICATION NUMBER: US/09/669,974
 CURRENT FILING DATE: 2000-09-26
 PRIOR APPLICATION NUMBER: US 09/377,155
 PRIOR FILING DATE: 1999-08-19
 PRIOR APPLICATION NUMBER: PCT/AU98/01031
 PRIOR FILING DATE: 1998-12-14
 PRIOR APPLICATION NUMBER: GB 9726398.2
 PRIOR FILING DATE: 1997-12-12
 NUMBER OF SEQ ID NOS: 33
 SOFTWARE: Patentin Ver. 2.0
 SEQ ID NO 32
 LENGTH: 1098
 TYPE: PRT
 ORGANISM: Haemophilus influenzae
 US-09-669-974-32

Query Match 5.5%; Score 117.5; DB 4; Length 1098;
 Best Local Similarity 19.1%; Pred. No. 0.04; Indels 229; Gaps 28;
 Matches 112; Conservative 75; Mismatches 170;

OY 7 TREKESADNONKOITDVSKISG-----LVNERKSEIIMAAKA--DANKHFGIN- 51
 DB 289 TVEFLSADTETITVYDSKENGKRETEVIGAKTSVIEKDGKLTGTGANKNETNVDGANA 348
 OY 52 -----MAIVTAGYVNDNSFNOSGWEAIOQALGTGETTSVDSST-----ALEBKYS 100
 DB 349 TEDADEBKGIVTAKDVI--DAVNKTGMRIKTTDANGONGDFATVAGSTNVTAFASGNGTTA 406

```

OY 101 SLANTNNKVVWVLSGFOHGD-----AFTRWLKI-----PENKO 132
DB 407 TYNCTDGTIVAYDAKVGGLKLDGDKIAADDTALTIVNGKNANNPKGVADVASTDEKK 466
OY 133 LFTENNIIT--IIGIDMTDT-----ENVIPGRYI-----NLTYKTEEAGW 170
DB 467 LVTAGGLVYALNSLMTWTYTAADGCTLDGNASEQEVKAGDKVTEKAGKMLVKOGGANF 526
OY 171 LAGVANASFLAKFPSPDKRSAYIGGGISPAVTF-----207
DB 527 TYSLDAL-----TGLTSTLGTGNNGAKTEINKDGLTTPANGAGANNANTISV 576
OY 208 -----IAGYLAGIKAN-----NL--KNSDKT 227
DB 577 TKDGISAGGQSVKNVNSGLKFKGDANFDPULTSSADNLTQONDAYKGLNLNDEKGDQOT 636
OY 228 KITTKKIEINLGFVDQDSTKERLEQIASKDKPSTLLAVAGPLTEIFSDIIANOND-RYL 286
DB 637 PVVADNTAATVG-DLRG-----LGMVISADK-----TTGGSTE-YHDOVRNANEVFKK 682
OY 287 IG-----VDTQSLVYTKT-----NKEF-----TSILKNLGYSVFVSLDL 323
DB 603 SGNGLNVSGKTYNGRREITFEELAKEGVKSNFTVKEKETSIVK-----VGDK 733
OY 324 YTKKSNRNLG--FEGKKSATVYLGKDFVDIADTS-----LEGN-----364
DB 724 YSKEDIDLTTGQPKLKDNTVAARYODKGVSVTDTEATITNKSGYVTGNQVADA 793
OY 365 -----DKKLAIEISEAKKEFEKTKTTPAEVVKTLTLEIPMPDK 404
DB 794 IAKSGFELGLADEA--DAKRAFDKTKALSA---GTEIYVAHDK 833

RESULT 13
US-09-268-347-44
: Sequence 44, Application US/09268347
: Patent No. 6335182
: GENERAL INFORMATION:
: APPLICANT: Loosmore, Sheena M.
: TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS
: FILE REFERENCE: 1038-860
: CURRENT APPLICATION NUMBER: US/09/268,347
: NUMBER OF SEQ ID NOS: 54
: SOFTWARE: PatentIn Ver. 2.0
: ID NO 44
: LENGTH: 1098
: TYPE: PRP
: ORGANISM: Haemophilus influenzae
US-09-268-347-44

```

Query Match 5.5%; Score 117.5; DB 4; Length 1098;
 Best Local Similarity 19.1%; Pred. No. 0.04; Indels 229; Gaps 28;
 Matches 112; Conservative 75; Mismatches 170;

```

OY 7 TREKESADNONKQITDVSIGS-----LVNERKSEIMAKA--DANKHFGILN- 51
DB 289 TVEFLSATETTTTIVYDREKCKRTEVAKIGAKTSYIKEDGKLFPGKANKETNKYDAGANA 348
OY 52 -----MAIYTAGTVNDNSFNOSGWEAIOQLGALTGETISVDSST---AELEGRYS 100
DB 349 TEDADEGKGLVYAKDVI--DAVNKTGMRIKTTDANGONGDFATVASGNVTFASGNGTJA 406
OY 101 SLANTNNKVVWVLSGFOHGD-----AFTRWLKI-----NLTYKTEEAGW 170
DB 407 TYNCTDGTIVAYDAKVGGLKLDGDKIAADDTALTIVNGKNANNPKGVADVASTDEKK 466
OY 133 LFTENNIIT--IIGIDMTDT-----ENVIPGRYI-----NLTYKTEEAGW 170
DB 467 LVTAGGLVYALNSLMTWTYTAADGCTLDGNASEQEVKAGDKVTEKAGKMLVKOGGANF 526

```

```

OY 171 LAGVANASFLAKFPSPDKRSAYIGGGISPAVTF-----207
DB 527 TYSLDAL-----TGLTSTLGTGNNGAKTEINKDGLTTPANGAGANNANTISV 576
OY 208 -----IAGYLAGIKAN-----NL--KNSDKT 227
DB 577 TKDGISAGGQSVKNVNSGLKFKGDANFDPULTSSADNLTQONDAYKGLNLNDEKGDQOT 636
OY 228 KITTKKIEINLGFVDQDSTKERLEQIASKDKPSTLLAVAGPLTEIFSDIIANOND-RYL 286
DB 637 PVVADNTAATVG-DLRG-----LGMVISADK-----TTGGSTE-YHDOVRNANEVFKK 682
OY 287 IG-----VDTQSLVYTKT-----NKEF-----TSILKNLGYSVFVSLDL 323
DB 603 SGNGLNVSGKTYNGRREITFEELAKEGVKSNFTVKEKETSIVK-----VGDK 733
OY 324 YTKKSNRNLG--FEGKKSATVYLGKDFVDIADTS-----LEGN-----364
DB 724 YSKEDIDLTTGQPKLKDNTVAARYODKGVSVTDTEATITNKSGYVTGNQVADA 793
OY 365 -----DKKLAIEISEAKKEFEKTKTTPAEVVKTLTLEIPMPDK 404
DB 794 IAKSGFELGLADEA--DAKRAFDKTKALSA---GTEIYVAHDK 833

```

```

RESULT 14
US-08-923-992A-4
: Sequence 4, Application US/08923992A
: Patent No. 6280738
: GENERAL INFORMATION:
: APPLICANT: Tai, Joseph Y.
: APPLICANT: Blake, Milan S.
: TITLE OF INVENTION: No. 6280738-19a Fc Binding Forms of the Group B
: NUMBER OF SEQUENCES: 34
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
: STREET: 1100 New York Avenue, N.W., Suite 600
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20005
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/923,992A
: FILING DATE: 05-SEP-1997
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/024,707
: FILING DATE: 06-SEP-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Esmond, Robert W.
: REGISTRATION NUMBER: 32,893
: REFERENCE/DOCKET NUMBER: 1438, 0140001/RWE
: TELEPHONE: (202) 371-2600
: TELEFAX: (202) 371-2540
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1104 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-923-992A-4

```

Query Match 5.4%; Score 116.5; DB 4; Length 1104;
 Best Local Similarity 17.9%; Pred. No. 0.051;
 Matches 85; Conservative 78; Mismatches 208; Indels 103; Gaps 13;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2002, 15:29:13 ; Search time 56.14 seconds
(without alignments)
892.310 Million cell updates/sec

Title: US-09-676-249A-2

Perfect score: 2299
Sequence: 1 MKRRKWKNFGLGLVPLS.....KQPDKQKISLTKLTIDINKN 451

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Minimum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2299	100.0	451	AAU01859	Mycoplasma hyopneum
2	2109	91.7	423	AAU01860	Mycoplasma hyopne
3	714.5	31.1	453	AAK40856	43rd regression as
4	694.5	30.2	432	AAK67582	Cancer metastasis
5	655	28.5	448	AAK93343	Gene encoding the
6	394.5	17.2	428	AAK05332	Inflamatory cytol
7	387.5	16.9	429	AAK22727	Membrane protein M
8	227	9.9	351	AAK1632	Streptococcus pneu
9	203	8.8	328	AAK55066	Streptococcus pneu
10	182	7.9	339	AAK00049	Enterococcus faeca
11	182	7.9	361	AAK00048	Enterococcus faeca

12	179	7.8	341	19	AAK61758
13	173	7.5	166	22	AAU07352
14	172	7.5	341	14	AAK33280
15	165.5	7.2	340	14	AAK31013
16	162	7.0	322	20	AAK19985
17	162	7.0	339	19	AAK61757
18	162	7.0	339	20	AAK19984
19	150	6.5	223	21	AAK91298
20	144	6.3	353	22	AAE03741
21	141.5	6.2	889	19	AAK68208
22	131.5	5.7	1786	18	AAK24790
23	129.5	5.6	1558	18	AAK18324
24	129	5.6	441	18	AAK28236
25	124.5	5.4	1135	21	AAK84460
26	122.5	5.3	2285	20	AAK98149
27	122.5	5.3	2346	22	ABB63519
28	122	5.3	2411	21	AAK23860
29	121.5	5.3	834	21	AAK93404
30	121.5	5.3	1099	19	AAK40538
31	120	5.2	719	22	AAK79241
32	119	5.2	776	22	ABB52597
33	119	5.2	892	19	AAK53313
34	117.5	5.1	1073	21	AAK01837
35	117.5	5.1	1079	21	AAK01836
36	117	5.1	2353	17	AAK93393
37	116.5	5.1	873	19	AAK68207
38	116	5.0	1300	22	AAU33407
39	116	5.0	1300	22	AAU35313
40	115.5	5.0	1164	17	AAK85781
41	115.5	5.0	1164	17	AAK40537
42	115.5	5.0	1164	21	AAK84459
43	114.5	5.0	1166	20	AAK08643
44	114	5.0	1333	22	AAU35343
45	113.5	4.9	466	16	AAK75739

ALIGNMENTS

RESULT 1	AAU01859	standard; Protein; 451 AA.
XX	AAU01859;	
AC	07-SEP-2001	(first entry)
XX		
DE	Mycoplasma hyopneumoniae MHP3 antigen.	
XX		
KW	MHP3; antigen; vaccine; enzootic mycoplasma pneumoniae;	
KW	antibody; immunoassay; immunotherapy; anti-idiotypic antibody.	
XX		
OS	Mycoplasma hyopneumoniae.	
XX		
FH	key	Location/Qualifiers
FT	Misc-difference 7	/note= "Encoded by TGA"
FT	Misc-difference 99	/note= "Encoded by TGA"
FT	Misc-difference 138	/note= "Encoded by TGA"
FT	Misc-difference 152	/note= "Encoded by TGA"
FT	Misc-difference 174	/note= "Encoded by TGA"
FT	Misc-difference 198	/note= "Encoded by TGA"
FT	Misc-difference 246	/note= "Encoded by TGA"
FT	Misc-difference 246	/note= "Encoded by TGA"
XX		
PN	EPI090995-A2.	
XX		
PD	11-APR-2001.	

B. burgdorferi ant
Interleukin-X (IL-
p39-beta. Borrel
p39-alpha. Borrel
B. burgdorferi ant
B. burgdorferi ant
B. burgdorferi ant
Group B Streptococ
Borrelia burgdorfe
M. catarrhalis str
P. falciparum live
Plasmodium falcipa
Amino acid sequenc
Bacillus subtilis
Drosophila melanog
Haemophilus influe
Amino acid sequenc
Mutant C-beta prot
Escherichia coli p
M. catarrhalis str
Haemophilus influe
Haemophilus influe
Haemophilus adhesi
M. catarrhalis str
Enterococcus faeca
Enterococcus faeca
Group B Streptococ
Group B Streptococ
Amino acid sequenc
S. aureus SdrE pro
Enterococcus faeca
B31 outer surface

XX 26-SEP-2000; 2000EP-0308421.
 XX
 XX 29-SEP-1999; 99US-0156602.
 XX
 XX (PRIZ) PRIZER PROD INC.
 XX
 XX King KW, Madura RA, Rosey EU;
 XX
 XX WPI: 2001-309781/33.
 DR N-PSDB; AAS03285.
 XX
 PT New apoprotein antigens encoded by mhp3 gene from Mycoplasma
 PT hyopneumoniae useful as a vaccine for treating or preventing diseases
 PT caused by Mycoplasma hyopneumoniae
 XX
 PS Claim 11; Page 17-18; 38pp; English.

XX The sequence is Mycoplasma hyopneumoniae MHP3 antigen. MHP3 antigen
 and its fragments are useful in manufacturing a vaccine for treating or
 preventing a disease or disorder in an animal, especially pig, caused
 by M. hyopneumoniae infection e.g. enzootic mycoplasma pneumonia.
 CC The mhp3-encoded proteins may be used as immunogens to generate
 CC antibodies which immunospecifically bind such an immunogen. The
 CC antibodies generated against the antigen are useful in diagnostic
 CC immunoassays, passive immunotherapy and generation of anti-idiotypic
 CC antibodies. Mhp3 proteins may also be used in immunoassays, e.g. to
 CC detect or measure in a biological sample from a vaccinated or
 CC potentially infected test animal the presence of antibodies to the
 CC antigen, and thus to monitor the immune response and/or to diagnose
 CC infection of the animal.
 XX
 XX Sequence 451 AA:

Query Match 100.0%; Score 2299; DB 22; Length 451;
 Best Local Similarity 100.0%; Pred. No. 5.2e-165;
 Matches 451; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKKIKMKNKFLGLVFPPLSAITATISAGCWDKETTKEEKSADNQNKOITDVSKISGLVNE 60
 DB 1 mkkkikmknkflglvfpplsaitatissagcwkdkctckeekssadnqnkqitdvskisglvne 60
 QY 61 RKSEMAAKADANKHFGGLMAIVTGGTVNDNSFNQSSWEAIIQOLGALTGGEITVSVDST 120
 DB 61 rkseimaaakadankhfgglmaivtggtvndnsfnqssweaiiqolgaltggetlsvdsst 120
 121 AELGKYSYSLANTNKNVWVLSGFQHDAPTRWLKIPENKOLFTEKNITIIIGIDWTDENV 180
 DB 121 aelgkysylantnknvwlsgfghdaptftrwlkipenkolfteknitiiigidwtdenv 180
 QY 181 IPTGRYINLTYKTEBAGMLAGYANASFLAKKPPSDPTRKSAIVIGGISPAVTDEIAGYL 240
 DB 181 iptgryinltyktebagmlagyanasflakkrpsdpkrsaiiviggispaavtdeiaagl 240
 QY 241 AGIKAMNIAKNSDKKRTTKTDKTEINLGFVDVDTSTKERLEQJASDKDSTLLAVAGPLTE 300
 DB 241 agikamniaknsdkkrttktdkteinlghfvdvdtstkerleqjaskdkdstllavagplte 300
 QY 301 IFSODIANONDRYLIGVTDOSLVYTKRNFPTSIKNLGYSVPSVSDLYTKRSNRN 360
 DB 301 ifsdianondrylignvtdoslvvtytkrnfptsiiknlgyssvpsvdsdlytkrsnsrn 360
 QY 361 LAGFEFGKSAIVYLGIDRFVDIADTSLGENDKRLATEAISAAKKEFEKTKTIPAEV 420
 DB 361 lagfefgkksaivylgidrfvdiadtstlegndkrlateaisaakefektktipaeav 420
 QY 421 RKTLEIPMPKQPPKQOESLDKLTIDINKN 451
 DB 421 rktleipmpkqppkqoesldkltidinkn 451

RESULT 2

AAU01860
 ID AAU01860 standard; Protein; 423 AA.
 XX
 XX AAU01860;
 AC
 XX
 XX 07-SEP-2001 (first entry)
 DT
 XX
 XX Mycoplasma hyopneumoniae recombinant MHP3 antigen.
 DE
 XX
 XX MHP3: antigen; vaccine; enzootic mycoplasma pneumonia; mutant;
 KM antibody; immunoassay; immunotherapy; anti-idiotypic antibody; muclein.
 XX
 XX Mycoplasma hyopneumoniae.
 OS
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 70
 FT Misc-difference 70
 FT /note= "Encoded by RCA"
 FT Misc-difference 417..422
 FT /note= "Encoded by AATWACGATAT"

XX EPI090995-A2.
 XX
 XX 11-APR-2001.
 PD
 XX
 XX 26-SEP-2000; 2000EP-0308421.
 PF
 XX
 XX 29-SEP-1999; 99US-0156602.
 PR
 XX
 XX (PRIZ) PRIZER PROD INC.
 PA
 XX
 XX King KW, Madura RA, Rosey EU;
 XX
 XX WPI: 2001-309781/33.
 DR N-PSDB; AAS03286.
 DR
 XX

PT New apoprotein antigens encoded by mhp3 gene from Mycoplasma
 PT hyopneumoniae useful as a vaccine for treating or preventing diseases
 PT caused by Mycoplasma hyopneumoniae
 PT
 XX
 XX Claim 4; Page 19-21; 38pp; English.

XX The sequence represents Mycoplasma hyopneumoniae a recombinant MHP3
 CC antigen lacking the first 28 amino acids (the putative signal sequence).
 CC MHP3 antigen and its fragments are useful in manufacturing a vaccine for
 CC treating or preventing a disease or disorder in an animal, especially
 CC pig, caused by M. hyopneumoniae infection e.g. enzootic mycoplasma
 CC pneumonia. The mhp3-encoded proteins may be used as immunogens to
 CC generate antibodies which immunospecifically bind such an immunogen. The
 CC antibodies generated against the antigen are useful in diagnostic
 CC immunoassays, passive immunotherapy and generation of anti-idiotypic
 CC antibodies. Mhp3 proteins may also be used in immunoassays, e.g. to
 CC detect or measure in a biological sample from a vaccinated or
 CC potentially infected test animal the presence of antibodies to the
 CC antigen, and thus to monitor the immune response and/or to diagnose
 CC infection of the animal.
 CC
 XX Sequence 423 AA:

Query Match 91.7%; Score 2109; DB 22; Length 423;
 Best Local Similarity 99.8%; Pred. No. 9.9e-151;
 Matches 414; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 30 WDKETTKEEKSADNQNKOITDVSKISGLVNRKSEIMAKADANKHFGGLMAIVTGGTV 89
 DB 2 wdkettkেকেসাদনকিতবস্কিসগ্লবনরকসৈমাকাদানকহফগলমাইবতগতব 61
 QY 90 NDNSFNQSSWEAIIQOLGALTGGEITVSVDSTPAELGKYSYSLANTNKNVWVLSGFQHDAP 149
 DB 62 ndnsfnqssweaiiqolgaltggetlsvdsstpaelgkysylantnknvwlsgfghdaf 121
 QY 150 TRWLKIPENKOLFTEKNITIIIGIDWTDENVIPTGRYINLTYKTEBAGMLAGYANASTLA 209
 DB 150 trwlkipenkolfteknitiiigidwtdenviptgryinltyktebagmlagyanastla 209

Db 122 twlikipenkgjfteknllilglwdtctenvjptgrylnlykteeagwlagyanasfia 181
 QY 210 KRPSPDTRSAIVIGGISPATVDFIAGVLAGIKAMNLKNSDKRTITTDKIEINLGF 269
 Db 182 kfpsdpktrsaiviggsipavtdfiagvlagikamnlknskckkclttcklelnlqf 241
 QY 270 VODTSTKERLEQIASKDKPSTLLAVAGPLTEIFSDIIANONDRYLIGVDTQSLVYTKR 329
 Db 242 vqdstckerleqaskdkpscllavagplteifsdliandndryliqvdtqslvyltkk 301
 QY 330 NKFETSLNLKGVSVSVSDLYTKSNSRNLAGFEFGKSAVYLGIDRFIDIDTSL 389
 Db 302 nfftsilknlygsvsvsdlytkksnrnlagfegkksatvlylqkdrfvdiadtsl 361
 QY 390 EGNDRKLTATEAISEAKKEFEKRTTPAEVRKLTLEIPEKPDKOPKQOESLTKL 444
 Db 362 egndkklataleaisaekfeektkltpaevrkltleipekpdqpkqeslkl 416
 RESULT 3
 AAR40856
 ID AAR40856 standard; Protein: 453 AA.
 AAR40856;
 XX 07-MAR-1994 (first entry)
 DE 43kd regression associated antigen.
 KW Regression associated antigen; tumour: immunotherapy;
 anti-idiotypic antibodies; antibodies; tumour regression.
 OS Mycoplasma hyorhinis.
 FH Key Location/Qualifiers
 FT Misc-difference 80 /note= "Tryptophan encoded by TGA, normal in
 Mycoplasma hyorhinis."
 FT Misc-difference 124 /note= "Tryptophan encoded by TGA, normal in
 Mycoplasma hyorhinis."
 FT Misc-difference 165 /note= "Tryptophan encoded by TGA, normal in
 Mycoplasma hyorhinis."
 FT Misc-difference 344 /note= "Tryptophan encoded by TGA, normal in
 Mycoplasma hyorhinis."
 PN US5242823-A.
 07-SEP-1993.
 PF 07-MAR-1986; 86US-0837494.
 PR 07-MAR-1986; 86US-0837494.
 PR 16-SEP-1987; 87US-0097910.
 PR 11-DEC-1987; 87US-0131815.
 PR 04-JAN-1988; 88US-0138923.
 PR 16-MAR-1990; 90US-0474730.
 PR 02-OCT-1992; 92US-0956546.
 XX (ITGE-) INT GENETIC ENG INC.
 PA Fared GC, Ghosh-dastidar P, Jar-how L, Sen A;
 PI WPI: 1993-295229/37.
 DR N-PSDB; AAQ47816.
 XX DNA encoding a regression-associated antigen from M. hyorhinis -
 PT is used to obtain prods. for diagnosis, localisation and therapy
 of tumours
 PS Disclosure; Figure 3; 40pp; English.

XX Regression associated antigens (RAA's) are identified in material
 CC from neoplastic cells by their immunological reactivity with
 CC regression associated antibodies from the serum of patients
 CC diagnosed as undergoing regression of a tumour. RAA's can be used
 CC for tumour immunotherapy and for producing and purifying antibodies
 CC which can be used for tumour diagnosis, localisation and therapy.
 CC The antibodies can also be used for the production of
 CC anti-idiotypic antibodies which can also be used in immunotherapy.
 SQ Sequence 453 AA:
 Query Match 31.1%; Score 714.5; DB 14; Length 453;
 Best local similarity 38.2%; Pred. No. 1.3e-45;
 Matches 175; Conservative 88; Mismatches 150; Indels 45; Gaps 16;
 QY 14 GLVPLSAIAT-ISAGCWDKETTKEKSADNQNKTITDVSKISGLVNERKSEIMAAKADA 72
 Db 12 gtlstvasvalfvscg----etdkegkll-----rifd-----nsfvkdrqaeleka---- 54
 QY 73 NKEFGIMALVTAGTYNDSEFNOSSWEAT---QQLGALTG-GETSVDSSTAELEGKY 127
 Db 55 -knidfnlvlltaggtvqdkstngslweavlehydqlektlnldrvsgetnngselgky 113
 QY 128 SSLANTNKNVWVLSGFQHDGDAFTRWLKIIPEN-----KQLFTEKNIIILGIDW--TDTEN 179
 Db 114 knflngnknwvlltgfgqgefekflqtdsngkkydsllaekvliavawdwlakedkd 173
 QY 180 VIPTRGYINLYTTEAGWLAGYANASFLAKFPSPDTRSAIVIGGISPATVDFIAGY 239
 Db 174 lkaghnfisllykteeagfiagyasckflaykfpndeakrtliapfggghagayctdfiaf 233
 QY 240 LAGIKAMNLKNSDKRTITTDKIEINLGFVODTSTKERLEQIASKDKPSTLLAVAGPLT 299
 Db 234 laglakyndnptakvclisdnlnldcgr-isdnkcatfinglvns--slvlpvagslt 290
 QY 300 EIRSDIIT--ANONDRYLIGVDTQSLVYTKKRFPTSLIKNLGYSVSVSDLYTKRSN 357
 Db 291 ssvvdalksknkdkthyligvtdqgsklfppa-lvfftslshkhgrltyevltdlwlkhd 349
 QY 358 SRNLAGEFEK---KSAIVYLGIKDRFVDIADTSLGNDKLTATEAISEAKKEFEKTK 413
 Db 350 skfigsfisfklnpanatvykgsdfigvsnstvavadakvkaqeflneatadffkqig 409
 QY 414 TTPAEVRKLTLEIPEM--PDKOPDKQOESLTKLITDIN 449
 Db 410 anpt-nyksvlgipmlindnakdnekalnelikkin 446
 RESULT 4
 AAR67582
 ID AAR67582 standard; Protein: 432 AA.
 AAR67582;
 AC 06-OCT-1995 (first entry)
 DT Cancer metastasis protein.
 XX Cancer metastasis protein.
 DE Cancer metastasis; clinical tests; detection.
 KW Homo sapiens.
 XX
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Misc-difference 55 /note= "corresponding codon TGA"
 FT Misc-difference 99 /note= "corresponding codon TGA"
 FT Misc-difference 140 /note= "corresponding codon TGA"
 FT Misc-difference 140 /note= "corresponding codon TGA"
 FT Misc-difference 319 /note= "corresponding codon TGA"

FT Misc-difference 420 /note= "corresponding codon TGA"
 FT Misc-difference 424 /note= "corresponding codon TGA"
 FT
 XX JP06319559-A.
 PN
 XX 22-NOV-1994.
 PD
 XX 12-MAY-1993; 93JP-0144165.
 PF
 XX 12-MAY-1993; 93JP-0144165.
 PR
 XX (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 PA
 DR WPI: 1995-040317/06.
 DR N-PSDB: AAQ79124.
 XX
 PT DNA encoding polypeptide involved in cancer metastasis - useful
 PT tests for the study of the mechanism of metastasis and in clinical
 tests
 PS
 XX Claim 1; Page 2; 13pp; Japanese.
 CC AAQ79124 encodes AAK67582 a protein involved in cancer metastasis,
 CC which may be used for the study of metastatic mechanisms, and for
 CC clinical tests to determine the presence or absence of cancer
 CC metastasis.
 CC
 XX
 SO Sequence 432 AA;

Query Match 30.2%; Score 694.5; DB 16; Length 432;
 Best Local Similarity 39.0%; Pred. No. 4e-44;
 Matches 168; Conservative 81; Mismatches 141; Indels 41; Gaps 14;
 OY 33 ETTKEKSAKDNDKOTDVSKISGLVNERKSEIMAAKADANKHFGUNMAIVTAGTVNDN 92
 DB 3 etdkegkll-----rlfd-----nsfvdkgaeleka-----knfdntvllagsgvqdk 48
 OY 93 SFNOSWEAI-----OOLGALTG-GETTSVDSSTAELEKYSLSLANRKNWVWLSGPHOD 147
 DB 49 sfngslweavlehydqlektlnldvsggetnngselgkyknfngknwvlltqgqgq 108
 OY 148 AETRMKIPEN-----KOLFTEKNIILIGIDW---TDENVIPTGRYIMLYKTEBAGWL 199
 DB 109 efpkflkqdsngkkyssdliaekvliavdwldskedkdlkaghflslllykteagfl 168
 OY 200 AGYANASFLAKKFPSPPTKRSALVIGGGISPAVTDPIAGYLAGIRAWMLKNSDKTKTTT 259
 DB 169 agyaasakflaykflpndeakrtlapfpg9hgagvtdflagflaglakyndnplqkvlsd 228
 OY 260 DEIEINIGFDVODSTKERLEQIASKDKRSTLAVAGPTEIFSDII--ANONDRYLIGV 317
 DB 229 mnlndtqf-lsndktactnglvnks--slvlpvgsitssvdaiksnxdktyllgv 285
 OY 318 DTDQSLVYTKTKNKEFTSILKNLGYSVFSVLSLDYTKRSNBNLAGFERGK---KSATV 373
 DB 286 dtdqsklifepa-tvfftslekhlgtrlyvltidwllkkesdflsfsfkltpnpanatv 344
 OY 374 YLGIDRFPVADIDTSLLEGNDKLALTEAISAEKKEFEKTKTTPAEVKTLEIPEK---P 430
 DB 345 ykgjaddfdvgsnsvadadkvakegelineatadfkkgiqanpt-nyxsvlgjlpmlnd 403
 OY 431 DKQPPKQOESL 441
 DB 404 ndakdnekasl 414

RESULT 5
 AAP93343
 ID AAP93343 standard; protein; 448 AA.
 XX

AC AAP93343;
 XX
 DT 01-AUG-1990 (first entry)
 XX
 DE Gene encoding the 43 kd regression-associated antigen (RA Ag) of
 DE Mycoplasma hyorhinis.
 DE
 KW Tumour regression-associated antigens (RA Ag); in vivo imaging;
 KW therapy monitoring; cancer therapy; Mycoplasma hyorhinis.
 XX
 OS Mycoplasma hyorhinis.
 XX
 XX
 FT Key Location/Qualifiers
 FT Peptide 1..25
 FT Protein 26..448
 XX
 PN EP308265-A.
 PD
 PD 22-MAR-1989.
 XX
 PF 16-SEP-1988; 88EP-0308625.
 XX
 PR 11-DEC-1987; 87US-0131815, US-097910.
 XX
 PA (ITGE-) INT GENETIC ENG INC.
 PI
 PI Fareded GC, Sen A, Ghosh-Dastidar P, Liu A, Lee JH;
 DR WPI: 1989-087638/12.
 DR N-PSDB: AAN90684.
 XX
 PT Tumour regression-associated antigens and antibodies -
 PT used in diagnostic tests, monitoring course of therapy and for
 PT therapy in cancer patients
 XX
 PS Disclosure; 56pp; English.
 XX
 CC Regression-associated antigens may be used in diagnostic tests, eg in
 CC vivo imaging, for monitoring the course of therapy or for therapeutic
 CC purposes, eg active immunisation protocols in cancer patients or drug
 CC delivery systems by binding the drug to monoclonal or monospecific
 CC polyclonal Ab showing specific immunoreactivity with the Ag.
 CC
 XX
 SO Sequence 448 AA;

Query Match 28.5%; Score 655; DB 10; Length 448;
 Best Local Similarity 37.4%; Pred. No. 4e-41;
 Matches 171; Conservative 89; Mismatches 149; Indels 48; Gaps 20;
 OY 14 GLVFPLSAIAT-ISAGCWDKETTEKESADNDKOTDVSKISGLVNERKSEIMAAKADA 72
 DB 12 glstsvaavalfvsg-----etdkegkll-----rlfd-----nsfvdkgaeleka----- 54
 OY 73 NKHFGUNMAIVTAGTVNDNSFNOSWEAI-----OOLGALTG-GETTSVDSSTAELEKGY 127
 DB 55 -knfdntvllagsgvtdkgsfngslweavlehydqlektlnldvsggetnngselgky 113
 OY 128 SSLANTNNWVWLSGFOGDAFTRKLRIPEN-----KOLFTEKNIILIGIDW---TDEN 179
 DB 114 knfngnknvnulltqgqgqepkflkqtdsngkkyssdliaekvliavawdwlskdkd 173
 OY 180 VIPTGRYINLYKTEEAGWLAGYANASFLAKKFPSPPTKRSALVIGGGISPAVTDPIAGY 239
 DB 174 ltkghfisllykceagfiagyaaskflaykflpndeakrtlapfpg9hgagvtdflagf 233
 OY 240 LAGIRAWMLKNSDKTKTTTOKIEINIGFDVODSTKERLEQIASKDKPSTLAVAGPRT 299
 DB 234 laglakyndnplakvltismnlndtqf-lsndktactnglvnks--slvlpvgsilt 290
 OY 300 EIFSDII--ANONDRYLIGVTDQSLVYTKTKNKEFTSILKNLGYSVFSVLSLDYTKRSN 357
 DB 291 ssvdaiksnkdktyllgvtdqsklifepa-tvffit-lekhlgtrlyevltidwllkkesd 348

OY 358 SRNLAGEFEK-----KSATVYLGIKREVDIADPTSLGNDKILATEBAISEAKKEFEKTK 413
 Db 349 s-flgsfrsfkltnaanatvykgsd-flgvsnstvadckvkagelnl-atadffkkqg 405
 OY 414 TTPAEVREKTLLEPEM-PKQPDKQOESLDKLTIDIN 449
 Db 406 ampt-nyksvlgipmlindnakhnekalnelikkin 441

RESULT 6
 ID AAY05332 standard; peptide; 428 AA.
 AC AAY05332;

DT 25-JUN-1999 (first entry)
 DE Inflammatory cytokine inducer.
 KW Inflammatory cytokine inducer; thrombocytopaenia; therapy.

XX Synthetic.

PN JP11089582-A.

PD 06-APR-1999.

PF 19-SEP-1997; 97JP-0273649.

PR 19-SEP-1997; 97JP-0273649.

PA (CHUS) CHUGAI PHARM CO LTD.

XX WPI: 1999-281057/24.

DR N-PSDB; AAX33847.

XX New inflammatory cytokine inducer gene and polypeptide - useful for treatment of thrombocytopaenia

PT Claim 3; Page 16-17; 22pp; Japanese.

CC This sequence represents the inflammatory cytokine of the
 CC invention. The inflammatory cytokine can be used in a drug, which is
 CC useful for the treatment of thrombocytopaenia.

XX Sequence 428 AA:

Query Match 17.2%; Score 394.5; DB 20; Length 428;
 Best Local Similarity 26.8%; Pred. No. 1.6e-21;

Matches 126; Conservative 101; Mismatches 174; Indels 69; Gaps 20;

OY 5 IKWNKRLGIVPLSAI-ATISAGCWDKETTKEKSADONKQITDVSKISGLVNERKS 63
 Db 1 mksskkilllgl-splaaalpavavscgnndes-----nlsfkedsklyttanngfq 52

OY 64 EIMAAKADANKHFLNMAIVTAGVNDNSFNQSSWEAIOQLGALGETTSVDSSTAEI 123
 Db 53 vkhneel-----lklpvyitdegkiddsfngsaftakalnkcgtclnnevps-snf 106

OY 124 EGRYSSLANNTNKNVWVLSGFQHDAPFTRMKIPENKQLTENKNIILGIDWDTENVIP 183
 Db 107 esaynasalagkhivlwgfkbgqskgy--ldahreelerngkijgidf-dleleykw 163

OY 184 GRVINTVYTEAGWLAGVANSFLAKRPPSDTKRSIAIVIGGISPAVTDPLAGIAGI 243
 Db 164 --tysqdnikesaftctgyalaswseq---deskrvasfgyvltlnegrfakg1 218

OY 244 KAWNLKNSDKTKRI-TTDRKEINILGFDVODTSTKERLEQIAS-----KDRPSTLL 292
 Db 219 lyyngqh--kskslyhtspvkidsqf-----tagekmtvlynnvls tpadvkyvnpvill 271

OY 293 AVAGPLTEIFSDIIANONDRYLIGVDTQSLVYTKTKNKFETSLKNLGYSFVSLSDI- 351
 Db 272 svagpat--fetvrlankqgyvlgvdsdgml--gqkdrlltevlkhikgaye1ldil 327
 OY 352 -----YTKSNSRNLAGFEFGKKSATVYLGIDREVDIADPTSLGNDKIL-ATEAI 401
 Db 328 lekeegykyvrvdkkdkwshfqtq-----kekvyvayenlfnsleeqakinnk1 379
 OY 402 SEAKKEFEKTKTTPAEVREKTLLEPEMPD-KQPDKQOESLDKLTIDIN 450
 Db 380 keaikmfke-----lpedfvkylnsdkalldgnkldhvsrleaisalnk 425

RESULT 7
 ID AAW22727 standard; Protein; 429 AA.
 AC AAW22727;

DT 26-FEB-1998 (first entry)

DE Membrane protein M161Ag.

XX M161Ag; membrane protein; leukaemia; treatment.

XX Homo sapiens.

OS Key location/Qualifiers

FT Modified-site 121 /note= "selenocysteine"

FT Modified-site 163 /note= "selenocysteine"

FT Modified-site 185 /note= "selenocysteine"

FT Modified-site 348 /note= "selenocysteine"

FT Modified-site 358 /note= "selenocysteine"

PN JP09157295-A.

PD 17-JUN-1997.

PF 05-DEC-1995; 95JP-0344504.

PR 05-DEC-1995; 95JP-0344504.

PA (KAGA-) KAGAKU GIYUTSU SHINKO JIGYODAN.

XX WPI: 1997-369470/34.

DR N-PSDB; AAT75133.

XX Membrane protein M161Ag - useful in development of leukaemia

PT Treating agent

PS Claim 1; Fig 1; 6pp; Japanese.

CC This sequence is a membrane protein designated M161Ag. Oligonucleotide
 CC probes for the cDNA were designed from putative N-terminal M161Ag peptide
 CC sequences isolated from P39 cells. M161Ag can be used in the development
 CC of a leukaemia treating agent.

XX Sequence 429 AA:

Query Match 16.9%; Score 387.5; DB 18; Length 429;
 Best Local Similarity 26.8%; Pred. No. 5.4e-21;

Matches 127; Conservative 96; Mismatches 174; Indels 77; Gaps 21;

OY 5 IKWNKRLGIVPLSAI-ATISAGCWDKETTKEKSADONKQIT-----DVSKISGLVN 59
 Db 1 mksskkilllgl-splaaalpavavsc-----annedsnlsfkedsklyttlna 48


```

RESULT 11
AAV00048
ID AAV00048 standard; Protein; 361 AA.
XX
XX AAV00048;
AC
XX 20-APR-1999 (first entry)
DT
XX Enterococcus faecalis protein EF021.
DE
XX Enterococcus faecalis; infection; vaccine; immune response; diagnosis;
KW detection; attenuation; antigenic.
XX
XX Enterococcus faecalis.
OS
XX
XX WO9850554-A2.
XX
XX 12-NOV-1998.
XX
XX 04-MAY-1998; 98WO-US08959.
XX
XX 14-NOV-1997; 97US-0066009.
XX 06-MAY-1997; 97US-0044031.
XX 16-MAY-1997; 97US-0046655.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Bailey C, Choi CH, Hromocky J A, Kunsch CA;
XX
XX WPI: 1999-070095/06.
XX DR N-PSDB; AAX20038.
XX
XX New isolated Enterococcus faecalis polynucleotides - used to develop
XX products for the detection of Enterococcus and for use in vaccines
XX for prevention or attenuation of Enterococcus infection
XX
XX Claim 9; Page 100; 301pp; English.
XX
XX The present sequence represents a protein isolated from
XX Enterococcus faecalis. The present invention describes genes, proteins
XX and antigenic polypeptides isolated from E. faecalis. The proteins can
XX be used in vaccines for preventing or attenuating an infection caused
XX by a member of the Enterococcus genus in an animal. They can also be
XX used for detecting Enterococcus antibodies in a sample. The nucleotide
XX sequences can be used for detecting Enterococcus nucleic acids.
XX Products from the present invention can also be used for screening
XX compounds to identify agonists and antagonists of E. faecalis protein
XX activity.
XX
XX Sequence 361 AA:

```

```

Query Match 7.9%; Score 182; DB 20; Length 361;
Best Local Similarity 25.4%; Pred. No. 1.3e-05;
Matches 104; Conservative 51; Mismatches 139; Indels 116; Gaps 20;
OY 61 RKSEIMAAKADANKHFGIMNAVTAAGTVNDNSFNQSSWEAIOQLGA---LTGGE----- 112
D 29 ktaesgggkygde---ahsaviiltgtgvdksfngsqwegqawgkeldpegsxkyay 84
OY 113 ITSVDSS--TAELEGGYSSSLANTKNVWVLSGFQHGDAFTRMKLPENKQLFTEKNIIIL 170
D 85 lgsnadaaytltldqvasakftltfgl---gyltkda-----lssaadandpctnfv11 134
OY 171 GIDWDTENVIPITGRVYNLTVYKTEBAGMLAGYANASPLAKRPPDPTKRSATVIGGISP 230
D 135 dddqldgkknv-----vacltrdneaaylaayaaa-----netktnkvgtvgygeeg 179
OY 231 AVTD-FIAGYLAKIKAWNLKNSDKTKTITDQKIEINLGFVDVDTSTKRELEQIASKDRPS 289
D 180 vldtftfggfeqgv-adaekelgkeltvdt-----kyasfaadpa 218
OY 290 TLAVAGPLTEITFSDI-----ANQDR-VLIGVDIDQSL--VYT 326

```

```

D 219 kykalaamyqngvdlilflhasgatqgvfgqaekdlnesgsgdkvwvlgvdrdqdadgky- 277
OY 327 KTK-----NKFFTSLKKNLGSVFSVLSDLTYTKSNSNNLWGFEGKRSATVYIGIKRFR 381
D 278 ktkgkednftclstlkvygtavdianraleck-----fpgenlhvy-glkdg 326
OY 362 VDIADTSLSEGNDRKLATEAISEAKKEFEKTKTTPAEVRRKTLIEPMD 431
D 327 vldtdgyl--ndkt-----keavktakdvlgsgdvkvpkpe 361

```

```

RESULT 12
AAW61758
ID AAW61758 standard; Protein; 341 AA.
XX
XX AAW61758;
AC
XX 09-SEP-1998 (first entry)
DT
XX B. burgdorferi antigenic protein p39 beta.
DE
XX Borrelia burgdorferi; antigenic protein; p39 alpha; p39 beta;
KW Lyme borreliosis; Lyme disease.
XX
XX Borrelia burgdorferi.
OS
XX
XX US5780041-A.
XX
XX 14-JUL-1998.
XX
XX 01-MAR-1995; 95US-0396957.
XX
XX 05-MAR-1991; 91US-0664731.
XX 05-MAR-1990; 90US-0487716.
XX 19-FEB-1993; 93US-0030245.
XX 01-MAR-1995; 95US-0396957.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Schwan TG, Simpson WJ;
XX
XX WPI: 1998-413001/35.
XX DR N-PSDB; AAV35685; AAV35687.
XX
XX New antigenic proteins from Borrelia burgdorferi useful in vaccines
XX - for raising antibodies or for diagnostic detection of specific
XX antibodies
XX
XX Example 6; Columns 31-34; 33pp; English.
XX
XX This represents the Borrelia burgdorferi antigenic protein p39 beta.
XX The antigenic protein p39 alpha is immunoreactive with mammalian Lyme
XX borreliosis serum. The Borrelia antigenic proteins p39 alpha and p39
XX beta are used in vaccines to protect against Lyme disease, as assay
XX reagents to detect specific antibodies in the serum (diagnostic of Lyme
XX disease), and to raise antibodies, either for diagnosis (by detecting
XX the corresponding antigen) or in screening agents for ability to inhibit
XX expression of the proteins.
XX
XX Sequence 341 AA:

```

```

Query Match 7.8%; Score 179; DB 19; Length 341;
Best Local Similarity 26.9%; Pred. No. 2e-05;
Matches 101; Conservative 59; Mismatches 123; Indels 92; Gaps 25;
OY 72 ANKHGELMAIVTAAGTVNDNSFNQSSWEAIOQLGALTGERTSVDSSTAELEGRYS-- 129
D 24 sskkikismlyv---dgvldksfnsaneallrlkkdtpenleefv--calsgyssyv 78
OY 130 --LANTKN-----VWVLSGFQHGDAFTRMKLPENKQLFTEKNII--ILGIDWDTENVIT 181

```


Db 79 sddlnkrngsdlw-lvgymltda--sllvsenpki--sygldpdygddvqipenll 133
 QY 182 PTERYINLYKTEEAGMLAGY--ANASFLAKFPSPDTRKSAIVIG--GGISPAVTD-FI 236
 Db 134 a-----vfrveggafllgylaakksfsgk-----lfgigmknigvdafr 174
 QY 237 AGYAGIKAMNLIKNSDKKRTITTDKIEINLCFDPVDTSTKERLEQIASK--DKPSTLLAV 294
 Db 175 yyesgakyaa-kdlelseynsfadvdg-----rtlsakmgsxgidvlfh 221
 QY 295 AGPLTFEI-FSDIANDND-RYLIGVDTDSLVYTKTKNKEFTSILKNLGYSVFSVLSDLX 352
 Db 222 aagllaglyvleaaknlgdygvygaddqds--ylapkn-fltsvknlgdalyllfge-y 277
 QY 353 TRKSNRNLAGFEFGKRSATVYLGIKDRFVDIADTSLEGNDKRLAEALSEAKKEFEKT 412
 Db 278 lknrv-----wegk-----vvgmgldgvyglpnan-----ef-eyl 310
 QY 413 KTIPEAEVKRTLEIP 427
 Db 311 kvlerkivnkelivp 325
 QY 13
 AAD07352
 ID AAD07352 standard; Protein; 166 AA.
 AC AAD07352;
 XX
 DT 03-JAN-2002 (first entry)
 DE Interleukin-X (IL-X).
 XX
 KM Interleukin-X; IL-X; autostimulatory factor; Epstein Barr virus; EBV;
 KW Lymphoblastoid; B cell.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 150 /note= "Encoded by TGA"
 FT Misc-difference 160 /note= "Encoded by TGA"
 FT
 XX
 PN US2001019713-A1.
 PD 06-SEP-2001.
 PF 16-JAN-2001; 2001US-0760541.
 XX
 PA 08-AUG-1994; 94US-0287387.
 PA 11-FEB-1993; 93US-0026132.
 XX
 PI (BLAZ/) BLAZAR B A.
 PA (WEBB/) WEBB A C.
 XX
 PI Blazar BA, Webb AC;
 XX
 DR WPI: 2001-570181/64.
 DR N-PSDB; AAS11798.
 XX
 XX New IL-X polypeptides useful e.g. for enhancing the growth of cells in
 PT vitro, for growing primary human B cell cultures to facilitate
 PT proliferation of these cell lines, as therapeutics or prophylactics, as
 PT molecular weight standards -
 XX
 PS Claim 1; Page 7; 16pp; English.
 CC The invention relates to a novel secreted autostimulatory factor,
 CC interleukin-X (IL-X), from an Epstein Barr virus (EBV)-carrying
 CC lymphoblastoid cell line. IL-X proteins can be used to enhance the growth
 CC of cells in vitro, or to grow primary human B cell cultures to facilitate
 CC proliferation of these cell lines, as therapeutics or prophylactics, as

CC molecular weight standards, as inert proteins in an assay, or in the
 CC detection of antibodies that are immunoreactive with IL-X. Nucleotide
 CC sequences encoding the peptides could be used to produce recombinant
 CC peptides or they could be used as probes or primers for diagnostic
 CC and/or analytical polymerase chain reaction (PCR) procedures, or as DNA
 CC or RNA sizing standards. The present sequence represents the amino acid
 CC sequence of IL-X.
 XX
 SQ Sequence 166 AA;
 QY
 Query Match 7.5%; Score 173; DB 22; Length 166;
 Best Local Similarity 27.3%; Pred. No. 2.1e-05;
 Matches 50; Conservative 37; Mismatches 56; Indels 40; Gaps 8;
 Db 225 GGGISPAVTDPIAGYLAGIKAMNLIKNSDKKRTKI-TTDKIEINLCFDPVDTSTKERLEQIA 283
 Db 2 gggafpgvtlfneigafkqillynqkh-kssklyhspvldsgf-----tagekmvtvl 54
 QY 284 S-----KDKPSTLLAVAGPLTFEISDIANDNDRYLIGVDTDSLVYTKRKNKP 333
 Db 55 nvlstpadvkynphvllsvagpat--fetvrlankgyvlgvdsqgml--qdkdrll 110
 QY 334 TSLIKNLGYSVFSVLSPL-----YTRKSNRNLAGFEFGKRSATVYLGIKDRFVD 383
 Db 111 tsylvkhikgavyetlidlilekeegykykpyvdkdkakwshfgtc-----kekwi 162
 QY 384 IAD 386
 Db 163 vae 165
 RESULT 14
 AAR33280
 ID AAR33280 standard; Protein; 341 AA.
 XX
 AC AAR33280;
 XX
 DT 17-DEC-2001 (updated)
 DT 11-MAY-1993 (first entry)
 XX
 DE P39-beta.
 XX
 KM Polymerase chain reaction; PCR; primer; open reading frame; ORF; P39;
 KW antigen; 39-alpha; 39-beta; serum; human; Lyme patient; p39-beta;
 KM p39-alpha; promoter; -10; -35; ribosomal binding site; RBS.
 XX
 OS Borrelia burgdorferi.
 XX
 PN USN7664731-N.
 XX
 PD 01-DEC-1992.
 XX
 PF 05-MAR-1990; 90US-0487716.
 XX
 PR 05-MAR-1990; 90US-0487716.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
 XX
 PI Schwan TG, Simpson WJ;
 XX
 DR WPI: 1993-017799/02.
 DR N-PSDB; AAQ34672.
 XX
 XX New Borrelia burgdorferi protein and DNA isolates - used in
 PT detection of Lyme borreliosis infection and in prodn. of
 PT antibodies and vaccines
 XX
 PS Disclosure; Page 36-39; 69pp; English.
 CC The sequences given in AAR31013 and AAR33280 represent the B.
 CC burgdorferi proteins P39-alpha and P39-beta. The DNA encoding these
 CC proteins was isolated by polymerase chain reaction (PCR) using the

Db 286 nyglkegvvgfv-----rnpkmisf-----elekeidnlssklnkeiivp----- 326
QY 434 PDRQESLDKLTID 447
: | : | :
Db 327 --snkesyekfike 338

Search completed: July 18, 2002, 15:32:18
Job time: 185 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2002, 15:29:13 ; Search time 32.42 Seconds
(without alignments)
1336.715 Million cell updates/sec

Title: US-09-676-249A-2
Perfect score: 2299
Sequence: 1 MKKKIKNNKFLGLVLPPLS.....KQPDKQGESLDKLTIDINKN 451

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Minimum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	554.5	24.1	461	2 B90555	ABC transporter xy
2	223	9.7	350	2 G95097	lipoprotein [impor
3	219	9.5	374	2 C97965	conserved hypotet
4	207	9.0	350	2 F86804	basic membrane pro
5	193.5	8.5	357	2 D96986	probable lipoprote
6	187.5	8.2	516	2 C82946	hypothetical prote
7	185.5	8.1	350	1 C70009	ABC transporter (1
8	182	7.9	353	1 H71340	membrane lipoprote
9	181.5	7.9	359	1 F72418	basic membrane pro
10	178	7.7	341	2 E70147	basic membrane lip
11	174	7.6	357	2 AH1610	CD4+ T cell-stimul
12	169.5	7.4	360	2 H70147	basic membrane pro
13	169	7.4	357	2 AD1248	CD4+ T cell-stimul
14	168.5	7.3	525	2 C82914	conserved hypotet
15	164	7.1	524	2 D82944	hypothetical membr
16	162	7.0	339	2 F70147	basic membrane pro
17	156.5	6.8	591	2 D64204	membrane lipoprote
18	147	6.4	353	2 G70147	basic membrane pro
19	143.5	6.2	384	2 I40867	hypothetical prote
20	141.5	6.2	539	2 D82886	conserved hypotet
21	141	6.1	379	2 H75318	membrane lipoprote
22	140.5	6.1	337	2 B97374	Deinococcus radiod
23	140.5	6.1	337	2 AH2591	membrane lipoprote
24	132	5.7	547	2 E29504	mercuryl(I) reduct
25	132	5.7	657	2 S73428	probable lipoprote
26	130	5.7	484	2 B90524	hypothetical prote
27	129.5	5.6	1140	2 S73786	hypothetical prote
28	129.5	5.6	1558	2 B71603	RESA-H3 antigen pr
29	127.5	5.5	763	2 A82863	hypothetical prote

30	127	5.5	349	2 F84246	hypothetical prote
31	126	5.5	326	2 G95857	hypothetical prote
32	125	5.4	1635	2 A10452	hemolysin [importe
33	124.5	5.4	556	2 H82301	peptide ABC transp
34	124.5	5.4	2269	2 T28677	thoptry protein -
35	123.5	5.4	322	2 F84236	ABC transporter (1
36	123.5	5.4	626	2 C25035	colicin Ia - Esche
37	123.5	5.4	1031	2 C81302	probable type I sl
38	123.5	5.3	1223	2 E88451	protein K10D2.1 (l
39	122.5	5.3	553	1 SMEBH1	flagellar hook-ass
40	122.5	5.3	553	2 AH0640	flagellar hook-ass
41	122.5	5.3	2285	2 T12796	probable transglyc
42	122.5	5.3	2346	2 T13829	tptr homolog - fru1
43	121	5.3	1546	2 G90603	lipoprotein [impor
44	119.5	5.2	350	2 F70139	exported protein (
45	119.5	5.2	627	2 A41609	dnak-type molecule

ALIGNMENTS

RESULT 1
B90555
ABC transporter xylose-binding lipoprotein [imported] - Mycoplasma pulmonis (strain U
C:Species: Mycoplasma pulmonis
C:Date: 24-May-2001 #sequence-revision 24-May-2001 #text-change 03-Aug-2001
C:Accession: B90555
R:Chamand, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Gallison, F.; Moszer,
Nucleic Acids Res. 29, 2145-2153, 2001
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma p
A:Reference number: A99512; MID:21267165; PMID:11353084
A:Accession: B90555
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-461 <R>
A:Cross-references: GB:A445566; PID:g14089760; PIDN:CAC13519.1; GSPDB:GN00153
A:Experimental source: strain UAB C71P
C:Genetics:
A:Gene: MYPU_3460
A:Genetic code: SGC3

Query Match 24.1%; Score 554.5; DB 2; Length 461;
Best Local Similarity 30.3%; Pred. No. 9.6e-27;
Matches 140; Conservative 88; Mismatches 175; Indels 59; Gaps 13;

QY	5	IKNNKFLGLVPLSLATISACQWCKETTKREKSDNONKQT--DVKISGLVNER- 61
DB	1	MKLNK--KLFSILPVALLALALPATF-----VSCANPNKTNLSLSSKRTDLSQKE 51
QY	62	-----KSEIMAKADANKHFLNMAIYAGTVNDNSFNOSWEPATIGALTGGEI 113
DB	52	VTEQKIVENKIKQASLETK-----VLLITADGNIDKSFNQGVYSQKTLDFVQKAY 106
QY	114	TS-----VDSSTAELEGKYSSLANTNNVWVLSGFQHDGAFTRMLKIPENK 159
DB	107	KSONKEAENOHKLDNYNSAVKXDEQNYKVALDRGYTWITLFGQOENETENFENDNNL 166
QY	160	QLTFEKIIILGIDWDPTEN-VIPTGRYINLYTTEBAGVLACANMSFLAKKPPSDPTK 218
DB	167	RRRENKRVKTIIGVWADNANSKIPQGSLSLFTTEBAGWQASADFLGTAYANNEAK 226
QY	219	RSALVIGGISPATVDFIAGYLAGIKMN--LKNSDKRTITTDKIRINLGFVDQDSTK 276
DB	227	RAISAFEGGDPAGVTDFLNGFPBGRIRANNEAEANANKKVIIVSNVLVDLQGF-IPNAEK 285
QY	277	ERLEQIASKDKPSTLLAVAGPLEIEESDI--AMQDRYLIGVDQSLVYTKNKKFF 333
DB	286	EVSNNVETGKSTSLPVAGPFGVVDVLRKDTSDREITVGVDTSLSFTWDSKRF 345
QY	334	TSILKNGYGVSVSLDLYT-----KKSNSNLNAGFEFGKSAIVYVIGINDRFVDA 385
DB	346	TSIVKNTAFVYQTLALTLTKDESVTLKRGNDKFLDS---NPKNVLVLRGLSAKFPVNT 402

OY 386 DTSLGNDKLTAEISEA--KKEEKTITPAEVRKTL 425
DB 403 KRVKESIKTQADTSIQKADKMNPNPNKKEKENTDLE 444

RESULT 2

695097
Lipoprotein [Imported] - Streptococcus pneumoniae (strain TIGR4)

C:Species: Streptococcus pneumoniae

C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001

C:Accession: G95097

R:RefSeq: H. Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Held

on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfel,

Science 293, 498-506, 2001

A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,

Title: Complete genome sequence of a virulent isolate of Streptococcus pneumoniae.

Accession: A95000; MUID:21357209; PMID:11463916

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-350 <KUP>

A:Cross-references: GB:AE005672; PIDN:AAK74976.1; PID:914972319; GSPDB:GN00164; TIGR:SP4

A:Experimental source: strain TIGR4

C:Genetics:

A:Gene: SP0845

Query Match 9.7%; Score 223; DB 2; Length 350;
Best Local Similarity 24.8%; Pred. No. 1.7e-06;

Matches 110; Conservative .63; Mismatches 148; Indels 122; Gaps 22;

OY 1 MKKIKMKNFLGLVPLSATATIS-AGCMDEKTEKESADNOKQITDVSKISGLVN 59
DB 1 MKK---QWGLGLV---AAVAGLAACGRSSRNASSSDVTK----- 39
OY 60 EKSEIMAAKADANKHFGSLNMAIVTAGTVDNSFNQSSWEAIQOLGA--LTGGEITSV 116
DB 40 -----AAIVTDGVDKSFQSDAMEGLQAMGKEHNLKDNNGFTY 79
OY 117 DSTAELE-----GKSSSLANTNNVWVLSFGDQDAFTRKLPENKQLTETKNI 167
DB 80 FOSTSEADYANNLQQAASYNLI-----GFALNNA-----VDAKKEHTDLY 125
OY 126 ILIGDWTDETVNIPITGRYINLTYTEAGWLAGYANASFLAKRPPDPTKSAIVIGG 227
DB 126 VLDIVINDQKRV-----ASVTFADNMSGYLAGYAAK-----TTTKQVGFVGG 170
OY 228 I-SPAVTDFIAGYLAGIRAMNLIKNSDKTKITTDKIEINLGFVDVDTSTKERLEQIASKD 286
DB 171 IESEVISRFEAGFKAGV-----ASVDPSIKVVDYAG--SFG-DAKKGKTIAAQAAGAD 223
OY 267 KPSTLLAAGPL-TFIFSDIIANQDR-----YLIGVDTQSL--VTT-----KTKNKF 333
DB 224 ---IYQVAGGTGACGVFAEASLNSRENEKRWVIGVDRDQAEAGKTSKDGKSNFVL 280
OY 334 TSILKNLGYSVFSLDYTTKKSNSRNLAGFEFGKKSATVYLGIDRFVDIADTSLSEGD 393
DB 261 VSTLKQVGTIVDI-----SNKAERGEFPFGQ---VIYSLDKQVDLAVTNLSEBG 329
OY 394 KKLATEAISEAKKEFEKTKTIP 416
DB 330 KK---AVEDAKAKITLDSVKVP 348

RESULT 3

C97965

conserved hypothetical protein spr0747 [Imported] - Streptococcus pneumoniae (strain R6)

C:Species: Streptococcus pneumoniae

C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001

C:Accession: C97965

R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Dehoff, B.S.; B

e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.

Y, P.; Sun, P.M.; Winkler, M.E.

J. Bacteriol. 183, 5709-5717, 2001

A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.

A>Title: Genome of the bacterium Streptococcus pneumoniae Strain R6.

A:Reference number: A97872; MUID:21429245; PMID:11544234

A:Accession: C97965

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-374 <KUP>

A:Cross-references: GB:AE007317; PIDN:AAK99551.1; PID:915458340; GSPDB:GN00174

C:Genetics:

A:Gene: spr0747

Query Match 9.5%; Score 219; DB 2; Length 374;
Best Local Similarity 24.6%; Pred. No. 3.3e-06;
Matches 109; Conservative .62; Mismatches 150; Indels 122; Gaps 21;

OY 1 MKKIKMKNFLGLVPLSATATIS-AGCMDEKTEKESADNOKQITDVSKISGLVN 59
DB 25 MKK---QWGLGLV---AAVAGLAACGRSSRNASSSDVTK----- 63
OY 60 EKSEIMAAKADANKHFGSLNMAIVTAGTVDNSFNQSSWEAIQOLGA--LTGGEITSV 116
DB 64 -----AAIVTDGVDKSFQSDAMEGLQAMGKEHNLKDNNGFTY 103
OY 117 DSTAELE-----GKSSSLANTNNVWVLSFGDQDAFTRKLPENKQLTETKNI 167
DB 104 FOSTSEADYANNLQQAASYNLI-----FGVGLNNAVEAEKKEHTDLY 149
OY 168 ILIGDWTDETVNIPITGRYINLTYTEAGWLAGYANASFLAKRPPDPTKSAIVIGG 227
DB 150 VLDIVINDQKRV-----ASVTFADNMSGYLAGYAAK-----TTTKQVGFVGG 194
OY 228 I-SPAVTDFIAGYLAGIRAMNLIKNSDKTKITTDKIEINLGFVDVDTSTKERLEQIASKD 286
DB 195 IESEVISRFEAGFKAGV-----ASVDPSIKVVDYAG--SFG-DAKKGKTIAAQAAGAD 247
OY 267 KPSTLLAAGPL-TFIFSDIIANQDR-----YLIGVDTQSL--VTT-----KTKNKF 333
DB 248 ---IYQVAGGTGACGVFAEASLNSRENEKRWVIGVDRDQAEAGKTSKDGKSNFVL 304
OY 334 TSILKNLGYSVFSLDYTTKKSNSRNLAGFEFGKKSATVYLGIDRFVDIADTSLSEGD 393
DB 305 VSTLKQVGTIVDI-----SNKAERGEFPFGQ---VIYSLDKQVDLAVTNLSEBG 353
OY 394 KKLATEAISEAKKEFEKTKTIP 416
DB 354 KK---AVEDAKAKITLDSVKVP 372

RESULT 4

F86804

basic membrane protein A [Imported] - Lactococcus lactis subsp. lactis (strain IL1403

C:Species: Lactococcus lactis subsp. lactis

C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001

C:Accession: F86804

R:Bohlin, A.; Winkler, P.; Manger, S.; Jallion, O.; Malame, K.; Weisenbach, J.; Eh

Genome Res. 11, 731-753, 2001

A>Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis

A:Reference number: A86625; MUID:21235186; PMID:11337471

A:Accession: F86804

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-350 <STO>

A:Cross-references: GB:AE005176; PID:912724428; PIDN:AAK05356.1; GSPDB:GN00146

A:Experimental source: strain IL1403

C:Genetics:

A:Gene: bmpA

Query Match 9.0%; Score 207; DB 2; Length 350;

Best Local Similarity 24.2%; Pred. No. 1.7e-05;
Matches 107; Conservative 62; Mismatches 156; Indels 118; Gaps 20.

Qy	1	MKKIKKNNKFLGGLVEPPSAITFTIS----	AGCMQRETTKEKSNADNQNQKOTTVDSKISG	56	
		:	:		
Db	1	MKKR-----	VIANSALALASVAVLACGRSHDAAGSK-----	32	
Qy	57	LYNERKSEIIMAAADANKHEGLMMAIVTAGGVYNDNFNSQSMWEAIIQALGALG----	G 111		
Db	33	-----AKTD-----	LKAAIVTEIGGVNDRSFNSQMEGLQISMGKNNLKRGTG	75	
Qy	112	EITSVOSTAELEGCKYSLSLANTKNVNVLSGQHDGAFTRMLKIENKQLPFEKIIILG	171		
		:	:		
Db	76	YTFPQSNASADYTTNTNSAQOQYKLLFCTGSLDQATSAAK-----	NN-----PKSEFVLD	129	
Qy	172	IDMTDINVTPTGRYINLTYTEAEAGWLAGVYANASELAKFPSPDPIKRSAT-VIGGIGSP	230		
		:	:		
Db	130	SVIKDQNV-----	ASAFEDNESAYLAGVAAK-----	ATCNKNGIGFQGMQSD	174
Qy	231	AVYDFIAGTYLAGIKANMLKNSDKTKITITTDKIEINLGPVODTSTKRLEQIADSKDRST	290		
		:	:		
Db	175	VTFEFGGYEAGAKSVN-----	PDIKVOYAGSFSFPAAGKGTIAAAMYAGAGDV	224	
		:	:		
Db	291	LIAVAGFL-EIFESDIIA-----	NONDR-VLIGVDPDQSLV--YT-----	KTRNKEFTSIL	337
		:	:		
Db	225	VYOCAGVGICGVSEAKALNSTKNEADKWWVIGVDDQDZETLKYKSKQCKDSNPLYVSTI	284		
		:	:		
Qy	338	KNIGYSVFSVLSDLTYKRSKNSRNLAG--	EFGKRSATVYLGIKDFRYDIADTSLEGNDK	395	
		:	:		
Db	285	KEVG-----	NVVKDIADTKDKGRPGGTIVTYLDKNGVNLG-----	DSANSEIK-----	330
Qy	396	LATPAISEAKKEFEKTKTTPAE	418		
		:	:		
Db	331	-----	DAVAKAKADIIDGKITVPSK	350	

RESULT 5
D96986
Probable lipoprotein, Med/BMP family [Imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
M:Accession: D96986
R:Nolling, J.; Berton, G.; Omelchenko, M.V.; Markarova, X.S.; zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A>Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cldo
A:Reference number: A96900; PMID:21359325; PMID:21359325
A:Accession: D96986
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-357 <KUR>
Cross-references: GB:NE001437; PIDN:AAK78679.1; PTD:g15023581; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
C:Gene: CAC0702

```

Query Match      8.5%; Score 195.5; DB 2; Length 357;
Best Local Similarity 22.3%; Pred. No. 8.7e-05;
Matches 100; Conservative 72; Mismatches 156; Indels 121; Gaps 20;

QY      1 MKKKIKNNKFLGIVFLPSAIAITTSAGCWMCKETTKEEKSADONKQITDVSKISGLVNE 60
      ||| : : : : : |||
Db      2 IKKKI-----IAILTYMIVAGLFAC-----SSTSSGSGNS 33

QY      61 RKSEIMAAKADANKHFEGLNMAIVTAGGVINDSNFSQSSWEALIQQLGALTGEITVSDET 120
      ||| : : : : : |||
Db      34 K-----DTKK---VKVGLSTDEGGLINDKFSFQAGDEGCIKKAKAEYSVDYKALESRK 81

QY      121 A-ELEGGKSSSLANTKKNNWVLSGFQHGDAF-TRMILKIPENKQLFEE-----KNIIIG 171
      ||| : : : : : |||
Db      82 KDDVQPNQLSINDNSDLVFGVGVQAMADLAITAKRYDPKKRAIIDDAYDKQPKNI--- 137

```

```

Qy 172 IDWIDTLEVIPTGYINLYKTEEAGLAGYANNSFLAKFPEDPpKRRAI-VIggGISP 230
Db 138 -----MSLVEKEDEGSFLMG-----VIAgKM-----ITNNIGVGRDOP 173
Qy 231 AVTDFIAGYLAGIKAWMLKNSDKKTKITTDKIEINLGFVDQDST-KERLEQIASDKPS 289
Db 174 LINKFSGYIAGAKTVN-----PNITVER-----NTYNDYSOSKGEVATSLYNGGCDI 223
Qy 290 TLLAVAGPFIETIEDIIANONDR-----YLGVVDGSLVYTKRKNFFSILKNLGSVF 345
Db 224 VYHAAGGAGIGVF-DVAKSLCDQCKDQWALGIVKDQAAGLPKADYILTSMSVRVIAIAY 282
Qy 346 SVLSDLYTKKSNSRNLGFEFGKKSATVYLGIKDRFDIADITSEGNCKKLATLEAISAK 405
Db 283 NTVDLVYKGRK-----FEGGKVES---PGLKEDDGVAPTS-----NKHVPEVLSTVD 328
Qy 406 K-----EFEEKRTTIAAEVR 421
Db 329 KYKKAIIIDGKIIVDPYDKAKOTFEKTDOIK 357

```

```

RESULT      6
C82946
hypothetical protein UU012 [imported] - Ureaplasma urealyticum
C:Species: Ureaplasma urealyticum
C:Date: 18-Aug-2000 #sequence.revision 20-Aug-2000 #text-change 20-Aug-2000
C:Accession: C82946
R:Glass, J.I.; Leikowitz, E.J.; Glass, J.S.; Helner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to Genbank, February 2000
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a
A:Reference number: A82870
A:Accession: C62946
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-516 <GLA>
A:Cross-references: GB:AF002100; GB:AF222894; NID:56898946; PIDN:AF30417.1; GSPDB:GN
A:Experimental source: serovar 3, biovar 1
C:Genetics:
A:Gene: UU012
A:Genetic code: SGC3

```

```

Query Match 8.2%: Score 187.5; DB 2; Length 516:
Best Local Similarity 23.3%: Pred. No. 0.000467
Matches 117; Conservative 67; Mismatches 167; Indels 151; Gaps 25.

QY 1 MKKKIKMKNFGLGLVFLPSLAITATISAGCQMDKETTKEEKSADNQNKKQITDVSIGSLVNE 60
Db 1 MKKLKKNVFLFIAGSVFALGITIIVATISOTQSTL----- 35

QY 61 RKSEIMAAKADANKHFGMLMALTVTAGGVINDNSFNQSSWEALIQULGALTGETTSDST 120
Db 36 -----NYSQFYW-----TSPTSD 49

QY 121 AELEGKYSSTANTNKNWVLSGFQHGDAFTRMLKIPENKQ--LEPTEK---NIITL--G 171
Db 50 DEGPQTKTKSNASDCKRALLMRYGH-----PELQNALVNDKRDPMILIALIDG 99

QY 172 IDWTDITENVIPGT--RYINLTYKTEAGMAGLAYANASFLAKK--FPSDPTKRSATVIIG 226
Db 100 YVNNQNKKEFYKGAQDRVADVEYFKVEDAAFLGIIAAYMLNSNQAVFGAD---NKLTMG 155

QY 227 --GISPA-VTDFIACYLAGIKRWNMLKNSDKKK---ITTDKIEINL-----GND 269
Db 156 YVGINAKNTTTLAEDFJAGVKNANKELDKNKIKQEGTQETKTKWINEQVYVASESSAGGQ 215

QY 270 VQDSTKEKLEQIADCKPSTLLAAGPLTEI-FESDIIANDQRY-LIGD-----TDOSL 323
Db 216 SDSANAKKIIQELIRKG-ADLLPALAIQVYGAIVAEALITVTHNNGVIGSDVDEINDAQI 274

QY 324 VYTKKKKKFFFSIL--KNLGVSVSVLS--DLYTKKSNRMLAGREGEKSAITYVLGI- 377
Db 275 --NKTQDFINPHLSGNNN-GVTRFSRTKRLDTAATIKLEENAIKESLSKESDKDOIVTIGE 331

```

[illegible]

```
Db      292 PFGGE-----TLTYGLDQDGIGISPSKQNLSDDVI--KAVDKKRKRRIIDGLEIPATEKE   343
OY          412 TKTIPE 418
           ||    ||
Db       344 LKTFFAE 350

RESULT      8
H71340
membrane lipoprotein tmpc precursor - syphilis spirochete
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C:date: 07-Aug-1998 #sequence revision 07-Aug-1998 #text_change 23-Jul-1999
C:Accession: H71340.A43595.S2956I
R:Frisser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gerson, J.; Knaik, N.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Uterback, T.; Meythey, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A:title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A:Reference number: A71250; MUID:98332770
A:Accession: H71340
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-353 <COL>
A:Cross-references: GB:AEO01211; GB:AEO00520; NID:g3322582; PIDN:AMC65302.1; PID:g3332
A:Experimental source: strain Nichols
R:Schnous, L.M.: van der Helde, H.G.J.: van Embden, J.D.A.
Infect. Immun. 59, 3536-3546, 1991
A>Title: Characterization of the 35-kilodalton Treponema pallidum subsp. pallidum receptor protein p35
A:Reference number: A43595; MUID:91372962
A:Accession: A43595
A:Molecule type: DNA
A:Residues: 1-10, A, I, 12-158, 'R', 160-353 <SCH>
A:Cross-references: GB:X57836; MID:g48838; PIDN:CAA40968.1; PID:g581809
A>Note: This protein is shown to incorporate palmitic acid
C:Genetics:
A:Gene: tmpr; TP0319
A:start codon: GTG
C:Superfamily: ABC transporter ynfN
C:Keywords: blocked amino end; lipoprotein; membrane protein; cholesterol bond
F:1-20/Domain: signal sequence *status predicted<SIG>
F:21-353/Product: membrane lipoprotein tmpc *status predicted<MNT>
F:21/Modified site: fatty acylated amino end (Cys) (in mature form) *status predicted
F:21/Binding site: snr-2,3-diacylglycerol (Cys) (covalent) *status predicted
```


QY 425 EIP 427
DB 343 VVP 345

RESULT 9
E72418
basic membrane protein - Thermotoga maritima (strain MSB8)

C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: F72418
R:Neilson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999

A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72200; MUID:99287316

A:Accession: F72418
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-359 <ARN>
A:Cross-references: GB:AE001696; GB:AE000512; NID:g4980582; PIDN:AAD35196.1; PID:g498059
Experimental source: strain MSB8

A:Gene: TM0102

C:Superfamily: ABC transporter ynfN

Query Match 7.9%; Score 181.5; DB 2; Length 359;
Best Local Similarity 23.9%; Pred. No. 0.00064;
Matches 94; Conservative 60; Mismatches 176; Indels 63; Gaps 16;

QY 61 RKSEIMAAKADANKHFGILNMAIYTAGGVNDNSFNQSSWEAIOQLALTGGEITSYDS-S 119

DB 2 RKFVLISLMIFPAVALFEGFVIWTDVGGLOKSPFNGSTAGITQAHDELGIEAKVIOSTE 61

QY 120 TAELEKYSLSLANTKNVWVLSGFQHGDAFTRLKIPENKOLFTEKNIIILIGDWTDE- 178

DB 62 OSDYIPNLSEKAEADLVAVGFMNTNDFKVAKOYPRD-----YFVGIDITPPEG 112

QY 179 NYIPGRYINLPKTEAGHLAGYANASFLAKKPPSPDKRSALVIGGGIS-PAYVDFEA 237

DB 113 QIIP--NVLTFTKEQEAFLVGYAAM-----TKGMVGVGVGIPPIPERFRY 161

QY 238 GYLAIKANKLNKNSDKRTKI---TTDKIEINLGFVDOTSTKERLEQIAK----- 282

DB 162 GYEGAKIKTYSVLHK-KNVILRGVTDPEDPKPKKDLAMSGFAEADLYVHASGAGNGV 220

QY 283 --ASAKRPSTLAVAGPLFEISDIANONDRLIGVDTDQSLVYTKNKKFTSLIKNL 340

DB 221 IEAAREKFEFA-LAGSDKLVDLDYVTTNGKGFPAIGVDMQDYM---APGAVLTSAMKRV 276

QY 341 GYSVFEVLDTLYTKNSRNRLAGEFEFGKKSATVYLGIKDFVLDIADJSLGDNKRLATEA 400

DB 277 DVASTYGVWAVE-----GTFEGGHR---VLGISDAVGI--SPKMTYKGLVPNRV 322

QY 401 ISE---AKKEFEKTKTIP-AEEVARKLEIPEM 429

DB 323 IAEILYLEKIMKRGCTLKVPETQEEIDAFVYPOI 355

RESULT 10
E70147
basic membrane lipoprotein B (bmbp) - Lyme disease spirochete

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 29-Sep-1999

C:Accession: E70147; I40280; I40242
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
D.; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venter, J.C.

A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.

A:Reference number: A70100; MUID:98065943
A:Accession: E70147
A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
A:Residues: 1-341 <LTP>

A:Cross-references: GB:AE001143; GB:AE000783; NID:g2688279; PIDN:AAC6758.1; PID:g268

A:Experimental source: strain B31
R:Simpson, W.J.; Cleplak, W.
FEMS Microbiol. Lett. 119, 381-388, 1994

A:Title: Nucleotide sequence and analysis of the gene in Borrelia burgdorferi encoding

A:Reference number: I40289; MUID:94327086

A:Accession: I40290

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 1-232, 'A', 234-317, 'V', 319-341 <RES>

A:Cross-references: GB:L24194; NID:g508420; PIDN:AAAT2407.1; PID:g508422

R:Ojamaa, C.; Davidson, B.E.; Saint Girons, I.; Old, I.G.
Microbiology 140, 2931-2940, 1994

A:Title: Conservation of gene arrangement and an unusual organization of RNA genes 1

A:Reference number: I40241; MUID:95111614

A:Accession: I40242

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 1-44, 'A', 46-179 <RE2>

A:Cross-references: GB:L35050; NID:g516591; PIDN:AAC41402.1; PID:g551744

C:Superfamily: basic membrane protein C

Query Match 7.7%; Score 178; DB 2; Length 341;
Best Local Similarity 26.7%; Pred. No. 0.00097;
Matches 100; Conservative 60; Mismatches 123; Indels 92; Gaps 25;

QY 72 ANKHFGLNMAIYAGGVNDNSFNQSSWEAIOQLALTGGEITSYDSSTAELEKYS-- 129

DB 24 SSKKIKISMV---DGVLDKSFNSNANEMALRLKDPFENIEVTS--CAISGVSSIV 78

QY 130 --LANTNKN---VWVLSGFQHGDAFTRLKIPENKOLFTEKNII--ILIGDWTDEYNT 181

DB 79 SDLDNLKRNQSDLIW-LVGYMLTDA--SLVYSENPKI--SYGIDPIYGDVDQIENLI 133

QY 182 PGRYINLPKTEAGHLAGY--ANASFLAKKPPSPDKRSALVIG--GGISRAVND-FI 236

DB 134 A-----VFRVQGAFLAGYIAKKSESGK-----IGTIGMKNINIDAFR 174

QY 237 AGYIAGIKANKLNKNSDKRTKITTTDKIEINLGFVDOTSTKERLEQIAK--DKPSTLAV 294

DB 175 YGESGAKYAN-KDIELISYNSFSVDYIG-----RTIASKNYSGIDYIHF 221

QY 295 AGPLFEI-FSDITANOND-RYLIGVDTDQSLVYTKNKKFTSLIKNLGYSFVSVDLY 352

DB 222 AAGIAGIGVIEFKAKNLGDGYVIGADQDS--YLAPEK-N-FITSVIRKIGDALVLTIGE-Y 277

QY 353 TKNSNSNLAGFEFGKKSATVYLGIKDFVLDIADJSLGDNKRLATEAISEAKPEPEKT 412

DB 278 IKNNNV-----WEGGK--VYVQGLRGVIGLPMAN-----EF-EYI 310

QY 413 KITPAEEVRKLEIP 427

DB 311 KYLERKINKKEIIVP 325

RESULT 11
AH1610
CD4+ T cell-stimulating antigen, lipoprotein [imported] - Listeria innocua (strain CI
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AH1610
R:Glaser, P.; Frangoul, L.; Buchliesser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Fsihl,
Science 294, 849-852, 2001
A:Authors: Kreft, U.; Kuhn, M.; Kunst, F.; Kurapkai, G.; Madueno, E.; Maltournam, A.;
ok, C.; Schluter, T.; Simoes, N.; Tlerez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla

A:Title: Comparative genomics of *Listeria* species.
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: AB1610
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-357 <GLA>
 A:Cross-references: GB:AL592022; PIDD:CA96656.1; PID:g16413898; GSPDB:GN00178
 A:Experimental source: Strain C11P11262
 C:Genetics:
 A:Gene: tcsA

Query Match 7.6%; Score 174; DB 2; Length 357;
 Best Local Similarity 24.2%; Pred. No. 0.0018;
 Matches 101; Conservative 62; Mismatches 137; Indels 118; Gaps 23;

55 SGLV-----NKRSEIIMAAKADANKHFGIINMAIVTAGVNVNDSFNQSSWEAIIQOLG 106
 16 SGVVLGACGSSDDKKS-----GDDKSSKDTVAMVITDGGVDRSFNQSAWEGIQKFG 69
 107 ALT---GGEITSVDSSTAELEGKYSLSANTKNVWVLS-----GFQHGDAFTRLMK 154
 70 KANDMEKGTGYNVLOSASEADYK-----TNLNTAVRSDDYDLIGIGYKLDKDAIEEVS 123
 155 -IPEKKOLETEKNIITLIGDMDTEVNIPTGRYINLTYTEAGMLAGYANASFLAKKP 213
 124 QPKKQFALVDTPT-----DSDNNVVSIG-----FDDNGSYLVGVVAGL----- 163
 214 SDPTKSAIVIGGISPATVD--FIAGYLAGIKAMNLKNSDKTKITTDKIEINLGFVQ-- 271
 164 -TTTNKNGVFGVKGAVIDREFEGFTAGVKA-----VNNNAQIDVQY 205
 272 --DTSTKERLEQIASKDKPS---TLAVAGPLTEIFSDIANQDR-----YLIGVDT 319
 206 ANDPAKADKGOOIASMSVSGVDVIFHAAGGTGNGVFAE--AKNLKKDPSRAVAVIGVDR 264
 265 DQWDEGKVTANDKDYNTVLTSEIKRVDAV---DDLAT-----RTKAGDPPGGTK--- 312
 320 DQ-----SLVYTKRKNKFTSILKNLGSVFSVLSDLTYTKNSRNLAG--FEFGKSKAT 372
 373 VYLGIKDRFVDIADTSLSEGNDKLATEAISEAKKEFEKTKTIPAEVYKLTLEIEMP 430
 313 IEYGLDKNAVGLSE-----HODNISKDVLAKE---EYKQKIVDD-----IKVEKP 357

RESULT 12

347
 A:Title: Membrane protein D (bmpD) homolog - Lyme disease spirochete
 C:Species: *Borrelia burgdorferi* (Lyme disease spirochete)
 C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 29-Sep-1999
 C:Accession: H70147
 R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; White
 son, D.; Peterson, J.; Kariavasa, A.R.; Quackenbush, J.; Salzberg, S.; Han-
 son, M.; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
 Nature 390, 580-586, 1997
 A:Authors: Smith, H.O.; Venter, J.C.
 A:Title: Genomic sequence of a Lyme disease spirochete, *Borrelia burgdorferi*.
 A:Reference number: A70100; MUID:98065943
 A:Accession: H70147
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-360 <KLE>
 A:Cross-references: GB:AE001144; GB:AE000783; NID:g2688291; PIDD:AA891505.1; PID:g268829
 A:Experimental source: strain B31
 C:Superfamily: basic membrane protein C

Query Match 7.4%; Score 169.5; DB 2; Length 360;
 Best Local Similarity 25.4%; Pred. No. 0.0035;
 Matches 88; Conservative 64; Mismatches 137; Indels 57; Gaps 20;

65 IMAAKADANKHFGIINMAIVTAGVNVNDSFNQSSWEAIIQOLGALTGRTTSDSSTABLE 124

DB 34 VACSSDDKSEAKTVSLI--VDGAFDDKGFENSSKAIKKLADININIEKASTGNSYL 92
 125 GYSSILANTNKN--VWVLSGFQHGDAFTRLKIPENKOLFTEKNIITLIGDMDTEVNIPT 183
 93 GDIANLEQNSNLINGI--GFRSLDIL--FQKASEVVS---NYAI--IEGYDEIQIPK 143
 184 GRYINLTYTEAGMLAGYANASFLAKKPPSDPTKSAIVIGGISPATVD--FIAGYLAG 242
 144 -MLNLSFRSEBEAFLAGY-----FASKASKTGKIGFVGKGVGLSEFMYEAG 193
 243 IKAMNLKNSDKTKITTDKI-----EINLGFVQDPTSTKERLEQIASKDKPSTLLAVAGPL 298
 194 AKYAN-----SNIKVVSQYVGTGDFGLG-----RSTASNM-----YRQGVDIIFAAAG-L 238
 299 TEIFSDIANQ--NDRYILGVDTDOSLVYTKTKNKFISILKNLGSVFSVLSDLTYTKKS 356
 239 SGIGVIEAKELGPRPHYIIGVQDQSYL--APNNVIVSAKKVQSLMYSL-----TKKY 290
 357 NSRNLAGFEFGKKSATVYLGK--DRFVDIADTSLSEGNDKLATEAI 401
 291 LETGV--LDGCK---TMEGLKEKDELGLVLENLKNYSSEIYNKSL 331

RESULT 13

AD1248
 CD4+ T cell-stimulating antigen, lipoprotein [Imported] - *Listeria monocytogenes* (str
 C:Species: *Listeria monocytogenes*
 C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
 C:Accession: AD1248
 R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec
 D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Fsihl,
 D.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madheno, E.; Maitouram, A.;
 Ok, C.; Schlueter, T.; Simoes, N.; Tiller, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla
 A:Title: Comparative genomics of *Listeria* species.
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: AD1248
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-357 <GLA>
 A:Cross-references: GB:NC_003210; PIDD:CA99466.1; PID:g16410817; GSPDB:GN00177
 A:Experimental source: strain EGD-e
 C:Genetics:
 A:Gene: tcsA

Query Match 7.4%; Score 169; DB 2; Length 357;
 Best Local Similarity 23.8%; Pred. No. 0.0037;
 Matches 98; Conservative 63; Mismatches 144; Indels 106; Gaps 22;

QY 54 ISGLVNERKSEIIMAAKADANKHFGIINMAIVTAGVNVNDSFNQSSWEAIIQOLGALT----- 109
 19 ILGACGSSDDKSSDDKSSDF--TVAMVITDGGVDRSFNQSAWEGIQKFGANDMEK 76
 110 GGEITSVDSSTAELEGKYSLSANTKNVWVLS-----GFQHGDAFTRLMK--IPENKQ 160
 77 GTDGYNYLOSASEADYK-----TNLNTAVRSDDYDLIGIGYKLDKDAIEEVSQKPKNOF 130
 161 LFTENKNIITLIGDMDTEVNIPTGRYINLTYTEAGMLAGYANASFLAKKPPSDPTKRS 220
 131 AIVDTT-----DSDNNVVSIG-----FKDNDGSLVGVVAGL-----TTKTN 168
 221 AIVIGGISPATVD--FIAGYLAGIKAMNLKNSDKTKITTDKIEINLGFVQ--DPTSK 276
 169 KVFPGVGVKGVIDREFEGFTAGVKA-----VNNNAQIDVQYANDRAKA 212
 277 ERLEQIASKDKPS---TLAVAGPLTEIFSDIANQDR-----YLIGVDTQ----- 321
 213 DKGOOIASMSVSGVDVIFHAAGGTGNGVFAE--AKNLKKDPSRAVAVIGVDRQWBGK 271
 322 -SLVYTKRKNKFTSILKNLGSVFSVLSDLTYTKNSRNLAG--FEFGKKSATVYLGITKD 379

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2002, 15:29:13 ; Search time 17.51 Seconds

(without alignments)
997,288 Million cell updates/sec

Title: US-09-676-249A-2

Perfect score: 2299
Sequence: 1 MKKIKWKKFLGLGLVEPLS.....KPPDKQGESLDKLTIDINKN 451

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Minimum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	185.5	8.1	350	1 YUFN_BACSU	005522 bacillus su
2	184	8.0	325	1 BMPA_BORGA	031357 borrelia ga
3	182	7.9	353	1 TMPC_TREPA	P29724 treponema p
4	179	7.8	341	1 BMPB_BORBU	045011 borrelia bu
5	174	7.6	357	1 TCSA_LISIN	Q92677 listeria in
6	170	7.4	341	1 BMPB_BORGA	031362 borrelia ga
7	169.5	7.4	341	1 BMPD_BORBU	044743 borrelia bu
8	169	7.4	357	1 TCSA_LISMO	048754 listeria mo
9	167	7.3	339	1 BMPA_BORAF	031380 borrelia af
10	162	7.0	339	1 BMPA_BORBU	045010 borrelia bu
11	156.5	6.8	591	1 Y040_MYCGE	P47286 mycoplasma
12	154	6.7	341	1 BMPB_BORAF	031384 borrelia af
13	147	6.4	351	1 BMPC_BORBU	050169 borrelia bu
14	139	6.0	498	1 PEDB_STRPY	099751 streptococ
15	132.5	5.8	981	1 SCA4_RICFE	Q9A137 rickettsia
16	132	5.7	547	1 MERA_STRAU	P08663 staphylococ
17	132	5.7	657	1 Y040_MYCPN	P75062 mycoplasma
18	129.5	5.6	1140	1 YD76_MYCPN	P75405 mycoplasma
19	123.5	5.4	626	1 CE1A_ECOLI	P06716 escherichia
20	122.5	5.3	552	1 FLCK_SALTY	P15932 salmonella
21	118	5.1	943	1 LBPA_NEITM	006379 neisseria m
22	117	5.1	655	1 YKDA_MYCCA	P45615 mycoplasma
23	116.5	5.1	4427	1 YKDA_MYCCA	005470 bacillus su
24	116	5.0	798	1 YC00_MYCPN	Q50288 mycoplasma
25	115.5	5.0	1164	1 BCGO_STRAG	P27951 streptococ
26	115.5	5.0	2022	1 ANP1_ONCYO	P21449 onchocerca
27	114.5	5.0	1111	1 KIR1_YEAST	P28742 saccharomyc
28	113.5	4.9	627	1 DNK1_GUTR	P29215 guillardia
29	113.5	4.9	1076	1 RPOB_ASTLO	P27059 astasia lon
30	112.5	4.9	944	1 LBPA_NEIMA	Q91K44 neisseria m
31	112	4.8	672	1 Y4A2_MYCPN	P75072 mycoplasma
32	111	4.8	848	1 Y4A7_SCHPO	Q09719 schizosacch
33	111	4.8	1902	1 P3P_LACLC	P15292 lactococcus

34	110.5	4.8	998	1 SCA4_RICAK	Q9A197 rickettsia
35	110	4.8	445	1 WAPA_STRMU	P11000 streptococ
36	110	4.8	626	1 PARC_BORBU	051066 borrelia bu
37	110	4.8	1957	1 YD86_SCHPO	Q10411 schizosacch
38	109.5	4.8	862	1 SLA2_BACAN	P94217 bacillus an
39	109.5	4.8	1176	1 SLAP_BACSH	P38537 bacillus sp
40	109.5	4.8	1290	1 XCPG_XENLA	P50532 xenopus lae
41	109.5	4.8	1701	1 MSPI_PLAFA	P13819 plasmodium
42	109.5	4.8	1701	1 MSPI_PLAFA	P08569 plasmodium
43	108.5	4.7	626	1 CE1B_ECOLI	P04479 escherichia
44	108.5	4.7	785	1 YG51_YEAST	P50089 saccharomyc
45	108	4.7	483	1 M6_STRPY	P08089 streptococ

ALIGNMENTS

RESULT 1	YUFN_BACSU	STANDARD;	PRT;	350 AA.
ID	YUFN_BACSU	005252;		
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Hypothetical lipoprotein yufN precursor.			
GN	YUFN.			
OS	Bacillus subtilis.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group;			
OC	Bacillus/Staphylococcus group; Bacillus.			
OX	NCBI_TaxID=1423;			
RN	SEQUENCE FROM N.A.			
RC	STRAIN=168;			
RA	Oudega B., Koningsheyn G., Duesterhoeft A., Hilbert H.;			
RL	Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.			
CC	-1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor (Potential).			
CC	-1- SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
DR	EMBL; Z93937; CAB07936.1; -			
DR	EMBL; Z99120; CAB15143.1; -			
DR	Subtilist; BG12349; yufN.			
DR	InterPro; IPR003760; Bmp.			
DR	Pfam; PF02608; Bmp. 1.			
DR	PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.			
KW	Hypothetical protein; Membrane; Lipoprotein; Signal;			
KW	Complete proteome.			
FT	STGNL 1 13 PROBABLE.			
FT	CHAIN 14 350 HYPOTHETICAL LIPOPROTEIN YUFN.			
FT	LIPID 14 14 N-ACYL DIGLYCERIDE (PROBABLE).			
SO	SEQUENCE 350 AA; 37349 MW; 16D5176A52A9284 CRC64;			

Query Match 8.1%; Score 185.5; DB 1; Length 350;
Best Local Similarity 24.8%; Pred. No. 0.00011;
Matches 106; Conservative 61; Mismatches 149; Indels 111; Gaps 21;

QY	21 AITATISGCCDKETKREKASADNONKOITDYKISGLVNERKSEIMAKADAKKHFGGLNM 80	
DB	6 AAOITILGACGNSR--KSSGSGEGKKNK-----FSV 32	
QY	81 AITTAGGTVDNDFNSQSWEAIOGLGLTG-----GEITSVDSSTAELEGKYSSTANTN 134	
DB	33 AMVITDVGGVDDKFNQSAWBGIOAFKGNGLKKGNKYDYLQSKSDADYTTNNKLAREN 92	

```

OY 135 KNWVLSQFHGDQFTPLRUKLPREKKOJLTETENIIILIGID-WTFQENVIPIGRYINLYTKT 193
Db 93 FDLIYGVCLMEDELS---ELADQ-----KNTNFALIIDAVYDKDNV-----ASTIFKE 138
OY 194 E6AGHLAGIYANASFLAKKFPSPDPKRS4I-VIGGISPAVTDFTAGIYLAGIKAMNKNSD 232
Db 139 QEGSFVLGVAAAL-----SSKSGKIGVGMEEBELKKFPEVGRQAOAVNPK-- 186
OY 253 KKT4ITTEDKIEINLGFVDQVDT-STKERLEQIASKDRPSTLAVAGPL-TEIFSDITAN- 308
Db 187 -----AVVEVAYAGGFQKADYQKAT4ESM-----YKSGVDVIYHAGATGIVGFTE-AKNL 236
OY 309 -----QNDRI4LGYDQDQ---SLVYTTKTKKNEFFSILAKNIGYSVFSVSLDLYTKKSNRN 360
Db 237 KKEPDKRDVWYVIGVDKDDQYAGQVEGTDDQNT4LTS4MK-----KQDVT4EDV-TEKASDGR 291
OY 361 LAGEEFGKRSATVYL4IKDREVDIADTSLSEGNDRK4LATEAISEAKKEF-----EEK 411
Db 292 FPGGB-----TLVYGLDQDQGVGISPSQ4N4SDVYI--KANVDKKKKRIIDGLEI4P4TEKE 343
OY 412 TKT4P4E 418
Db 344 LKTFRAE 350

```

RESULT						2
ID	BMPA_BORGA	STANDARD:	PRT:	325 AA.		
AC	OJ3357; OJ1360;					
DT	15-DEC-1998 (Rel. 37, Created)					
DT	15-DEC-1998 (Rel. 37, Last sequence update)					
DT	15-DEC-1998 (Rel. 37, Last annotation update)					
DE	Basic membrane protein A precursor (Immunodominant antigen P39) (Fragment).					
GN	BMPA.					
OS	Borrelia garinii.					
OC	Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.					
OX	NCBI_TaxID=29519;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=PRI. AND PLT.					
RX	MEDLINE=96010210; PubMed=9350727;					
RA	Roesler D., Hauser U., Wilske B.;					
RT	"Heterogeneity of Bmpa (P39) among European isolates of Borrelia burgdorferi sensu lato and influence of interspecies variability on serodiagnosis.";					
CC	J. Clin. Microbiol. 35:2752-2758(1997).					
CC	-1 FUNCTION: NOT KNOWN; IMMUNOGENIC PROTEIN.					
CC	-1 SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor (Probable).					
CC	-1 SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS. (Probable).					
CC	-----					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).					
CC	-----					
DR	EMBL; X97244; CAA65883.1; -.					
DR	EMBL; X97238; CAA65877.1; -.					
DR	InterPro; IPR003760; Bmp.					
DR	InterPro; IPR000437; Prok_Lipoprot.					
DR	Pfam; PF02608; Bmp; 1.					
DR	PROSITE; PS00013; PROKAR_LIPOPROTEIN_PARTIAL.					
KW	Antigen; Membrane; Lipoprotein; Signal.					
FT	NON_TER	1				
FT	SIGNAL	<1	3		PROBABLE.	
FT	CHAIN	4	325		BASIC MEMBRANE PROTEIN A.	
FT	LIPID	8	4		N-ACTYL DIGLYCERIDE (PROBABLE).	
FT	VARIANT	4	8		G -> D (IN STRAIN PLT.)	
FT	VARIANT	12	12		S -> N (IN STRAIN PLT.).	

FT	93	93	P -> S (IN STRAIN PLY) .
FT	VARIANT	111	T -> A (IN STRAIN PLY) .
FT	VARIANT	180	N -> D (IN STRAIN PLY) .
FT	VARIANT	205	S -> G (IN STRAIN PLY) .
FT	VARIANT	239	I -> V (IN STRAIN PLY) .
FT	VARIANT	243	S -> S (IN STRAIN PLY) .
FT	VARIANT	250	S -> A (IN STRAIN PLY) .
FT	VARIANT	257	N -> S (IN STRAIN PLY) .
FO	SEQUENCE	325480 MM; D19288JL2AAB1C158 CRC644	

Query Match	8.0%;	Score 184;	DB 1;	Length 325;
Best Local Similarity	26.1%;	Pred. No. 0.00012;		
Matches	93;	Conservative	56;	Mismatches 135; Indels 72; Gaps 18

QY	77	GLNNAIYLAGTAVNDNFENNQSNMEALIQOALGALGETTSVDSST-----NELGKXSLAN	132
Db	13	GIPVSVIYVNGTETPDKKFNSALNGIKAKKEEKIEVLKSTNSYSLDLEG-----LKD	68
QY	133	TNRNVVWLSGFQHD-AFTFMLKIPENKOLFTEKNIILIGIDMTDVENYIPGRYINILTY	191
Db	69	AGSNLIMLIGRFSVDVAKAVSLQNEPKYAI-----IDPVSESIPT-NLVGMTE	118
QY	192	KTEAGMLAGYANASFLAKFPSPDPTRKSAIYVGGISPAVYD-PLAGYIACIAKMNILKN	250
Db	119	RAQGAFLTGTYIAKV-----SKTGKIGPLGIGEEIYDAFPGYEGACAKAYAN--	166
QY	251	SDRKTKITTKTI-----EINLGFVDVQPTSTKERLEQIASDKDSTLAVAGPLTEIFSDII	306
Db	167	--KIKITSARIHSGFNDVEAG---RSVATKRMYSQGI-----DIHHAAISLGGIGALEV	214
QY	307	ANQ--NDRYILGVGDVQDSLYVYTKTKNKFPTSLIKNLGYSVFSYSLDYLYTKKSNRNLAGF	364
Db	215	AKELGSGHYITIGVDEQDSYL--APNNIITISAKKQVGRSL-NIFTSNLYLKTMT-----F	264
QY	365	EFGKSAATVYLGIRDFVVDIADTSLSEGNCKILATEAISAKKEEFEKTKFTIPAEVY	420
Db	265	EGGR---LINVGLKEGVYGVFK-----NPKMIP-----FELEKEIDMLSSKIINOEI	308

RESULT 3

ID	TMPC_TREPA	STANDARD;	PRT;	353 AA.
AC	P29724; O83339;			
DT	01-APR-1993 (Rel. 25, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Membrane lipoprotein tmpr precursor (Membrane protein C) (35 kDa antigen) (Lipoprotein Tpm35).			
GN	TMPC OR TP0319.			
OS	Treponema pallidum.			
OC	Bacteria; Spirochaetales; Spirochaetaceae; Treponema.			
OX	NCBI_TaxID:160;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-NICHOLS.			
RX	MEDLINE-91372962; PubMed-1894360;			
RT	Schouls L.M., van der Heide H.G.J., van Embden J.D.A.:			
RT	"Characterization of the 35-kilodalton Treponema pallidum subsp.			
RT	pallidum recombinant lipoprotein tmpr and antibody response to			
RT	lipidated and nonlipidated T. pallidum antigens."			
RL	Infect. Immun. 59:3536-3546(1991).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-NICHOLS.			
RX	MEDLINE-98332770; PubMed-9655876;			
RA	Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,			
RA	Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,			
RA	Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,			
RA	Khalak H., Richardson D., Howell J.K., Childsbarham M., Ufferback T.,			
RA	McDonald L., Attiach P., Bowman C., Cotton M.D., Fujii C., Garland S.,			
RA	Hecht B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,			
RA	Velter J.C.,			

```

RT *Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete."
RL Science 281:375-388(1998).
CC -I- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC (Probable).
CC -I- SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X57836; CAA40968.1; -.
DR EMBL: AB001211; AAC65302.1; -.
DR PIR: A43595; A43595.
DR PIR: S29561; S29561.
DR TIGR: TP0319; -.
DR InterPro: IPR003760; Bmp.
DR Pfam: PF02608; Bmp. 1.
DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
DR Antigen: Membrane; Lipoprotein; Signal; Complete proteome.
FT SIGNAL 1 20
FT CHAIN 1 353
FT LIPID 21 21
FT MUTAGEN 21 21
FT FT
FT CONFLICT 11 11
FT CONFLICT 159 159
FT SEQUENCE 353 AA; 37769 MW; CB373138C8C4337D CRC64;

Query Match 7.9%; Score 182; DB 1; Length 353;
Best Local Similarity 23.7%; Pred. No. 0.00018;
Matches 86; Conservative 69; Mismatches 134; Indels 74; Gaps 17;

OY 80 MAIVTGVNNSPNSQSSWEAIOQLGALTGRITVDSST-ALEIKYPSILANTNNKW 138
DB 42 VGMVTDGSDIDKSFNQWEGISRAQENNAKCKRYTASTDAEYVPSLSAFDENMGIV 101
OY 139 VLSGFGHDAFTRMKIPENKQLFTEKNIIIGIDMTDENVIPTGNVYINLYKTEAGW 198
DB 102 VACGSFLVFA-----VIFTSARFPKQKFLVIDAVVQDRNV-----VSAVGQNGESF 149
OY 199 LAGYANASPLAKKFPSPDKRSAL--VIG--GISPAVDFIAGYLAGIKANMLKNSDKR 254
DB 150 LVGVA-AALKAKE---AGKSAVGFIVGMEIGMPL--FEAGFEGVAKAVD----- 193
OY 255 TRITTDKIEINLGFVDYDT-STKERLEQJASK---DKPSTLLAVAG-----PLTEIFSDII 306
DB 194 -----PIQYVVEVANTFSDPKQALAAKIYDGVNVIFPVAGTGTGNGVKEARDR 246
OY 307 ANQNDRYLLIGVDTDSL--VYTKTKNKFSTSLKNGYSVFSVLSDIYTKKSSRMLAG 364
DB 247 LMGODVWVIGVBDQYMDGYDGSKSVLTSWKRA-----DVAARHISMAVDGSPG- 300
OY 365 EGGKSAIYVLGKDRFVADIATSLGNDKALATEAISEKKKFEFEKTKTTPAEVAKTL 424
DB 301 -----GQSIIMFGLEKRAVGIE-----ENPNLSAVMERKIRSFEEKI-----VSKETI 342
OY 425 EIP 427
DB 343 VVP 345

RESULT 4
BMPB_BORBU STANDARD: PRT: 341 AA.
AC Q45011; Q07954; Q31317; Q50168; Q44858;
DT 15-DEC-1998 (Rel. 37, Created)
DR 15-DEC-1998 (Rel. 37, Last sequence update)

```

```

DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Basic membrane protein B precursor.
GN BMPB OR B80382.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_Taxid=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SH-2-82;
RX MEDLINE=94327086; PubMed=8050720;
RT Simpson W.J., Cieplak W., Schrumpt M.E., Barbour A.G., Schwan T.G.;
RT Nucleotide sequence and analysis of the gene in Borrelia burgdorferi
RL encoding the immunogenic P39 antigen.
RL FEMS Microbiol. Lett. 119:381-388(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-297;
RX MEDLINE=97132632; PubMed=8978084;
RA Aron L., Toth C., Godfrey H.P., Cabello F.C.;
RT Identification and mapping of a chromosomal gene cluster of Borrelia
RT burgdorferi containing genes expressed in vivo.
RL FEMS Microbiol. Lett. 145:309-314(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 35210 / B31;
RX MEDLINE=98055943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.R., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Uitterback T., Wachtel L., McDonald L., Aitchak P., Bowman C.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia
RT burgdorferi."
RL Nature 390:580-586(1997).
RN [5]
RP SEQUENCE OF 1-179 FROM N.A.
RC STRAIN-212;
RX MEDLINE=9511614; PubMed=7812434;
RA Ojalmt C., Davidson B.E., Saint-Girons I., Old I.G.;
RT "Conservation of gene arrangement and an unusual organization of rRNA
RT genes in the linear chromosomes of the Lyme disease spirochaetes
RL Borrelia burgdorferi, B. garinii and B. alzellii."
RL Microbiology 140:2931-2940(1994).
CC -I- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC (Probable).
CC -I- SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L24194; AAA72407.1; -.
DR EMBL: U49938; AAC44713.1; -.
DR EMBL: X81517; CAA57237.1; -.
DR EMBL: AB001143; AAC66758.1; -.
DR EMBL: L35050; AAC41402.1; -.
DR TIGR: B80382; -.

```



```

RP SEQUENCE FROM N.A.
RC STRAIN-PBI;
RX MEDLINE-98010210; PubMed-9350727;
RA Roesler D., Hauser U., Wilske B.;
RT "Heterogeneity of BmpA (P39) among European isolates of Borrelia burgdorferi sensu lato and influence of interspecies variability on serological diagnosis."
RL J. Clin. Microbiol. 35:2752-2758(1997).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor (Probable).
CC -1- SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC
CC EMBL; X81518; CAA57238.1; -
DR InterPro: IPR003760; Bmp.
DR Pfam: PF02608; Bmp.1.
KW PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Membrane; Lipoprotein; Signal.
FT SIGNAL 1 14 PROBABLE.
FT CHAIN 15 341 BASIC MEMBRANE PROTEIN B.
FT LIPID 15 15 N-ACYL DIGLYCERIDE (PROBABLE).
SQ SEQUENCE 341 AA; 37236 MW; 442BEF0BEDDC9A CRC64;

Query Match 7.4%; Score 170; DB 1; Length 341;
Best local Similarity 26.8%; Pred. No. 0.00956;
Matches 91; Conservative 56; Mismatches 116; Indels 76; Gaps 20;

QY 87 GYVNDNSFNQSWEAIDQALGGETSVDSSTAELEGY----LANTNN-----VW 138
DB 36 GVDDDSFNSANRALRLREDEPENIEKVFSSNA--SGVSSVSDLDLNMKNGSDLIW 93
QY 139 VLSGFQHDGAFTRMLKPEKKOLFETKNIILIG-IDMTDVENYPRGRYINILPYKTEAG 197
DB 94 -LVGYMITDA-----SLVSLNPKISTYGIIDPVYDDVOIPKNLIGVFRIDQGA 143
QY 198 WLAGYANASFLAKRFPSPDKRSALVIGGISPAYTD-FLAGLAGIKANNLNKSDKRTK 256
DB 144 FLAGY-----IAKKSVS-----GKIFIGVKGDIYDAFRYGEAG-----AKYADGIE 189
QY 257 ITTDK-----IEINLGFVQDTSTKRELEQIASK--DKPSTLAVAGPLTEIFSDITANQN 310
DB 190 IVEYSNSFSFSDVNI-----RAIANKKYAKGIDIIHFMAGLAGI--GVIEAPK 235
QY 311 D-----RYLIGVDIPOSIVYRTKTKKPFSTSLKNGYVSFVSLDLYKTKNSRLAAGFEF 366
DB 236 ELGEGYIVADODQSHL---APRNFTTSYIAKNVGDALVLTSE-YLKNNT-----WEG 286
QY 367 GKRSATVYLGIKRFDVIADTS---LEGNDKRLATEAI 401
DB 287 GK---IIOGLRDGVGLSNANKEFYIKVIERKIVNEEI 322

RESULT 7
BMPD_BORBU STANDARD; PRT; 341 AA.
AC 044743;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Basic membrane protein D precursor.
GN BMPD OR BB0385.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN-JD1;
RX MEDLINE-96178617; PubMed-8606088;
RA Ramamoorthy R., Povinelli L., Philipp M.T.;
RT "Molecular characterization, genomic arrangement, and expression of bmpD, a new member of the bmp class of genes encoding membrane proteins of Borrelia burgdorferi."
RL Infect. Immun. 64:1259-1264(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 35210 / B31;
RX MEDLINE-98065943; PubMed-9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Karpavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Ullrich T., White L., McDonald L., Artlisch P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi."
RL Nature 380:580-586(1997).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor (Probable).
CC -1- SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC
CC EMBL; U35450; AAC43984.1; -
DR EMBL; AE001144; BAB91505.1; -
DR TIGR; BB0385; -
DR InterPro: IPR003760; Bmp.
DR Pfam: PF02608; Bmp.1.
KW PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Membrane; Lipoprotein; Signal; Complete proteome.
FT SIGNAL 1 16 PROBABLE.
FT CHAIN 17 341 BASIC MEMBRANE PROTEIN D.
FT LIPID 17 17 N-ACYL DIGLYCERIDE (PROBABLE).
SQ SEQUENCE 341 AA; 37163 MW; 63FE638FE81A91D5 CRC64;

Query Match 7.4%; Score 169.5; DB 1; Length 341;
Best local Similarity 25.4%; Pred. No. 0.001;
Matches 88; Conservative 64; Mismatches 137; Indels 57; Gaps 20;

QY 65 IMAKAAKHFGLMAIVAGTVNDNSFNQSWEAIDQALGGETSVDSSTAELE 124
DB 15 VACSSDDGKSEAKTVSLI-VDGAFDDKGFENSSKAIKRLKADLNINIEKASTGNSYL 73
QY 125 KYSSILANTNNK-NVWISGFQHDGAFTRMLKPEKKOLFETKNIILIGIMTQDENVIPT 183
DB 74 GDIANLEDGNSNLIMGI-GRLSDIL--PQASRNSV-----NVAI--IEGVDEIDIPK 124
QY 184 GRYINLTYKTEAGWLAGYANASFLAKRFPSPDKRSALVIGGISPAYTD-FLAGLAG 242
DB 125 -NLNIFSRSEVAFVLAG-----FASKASKTGKIGFVGGRVKVLESFMYGEAG 174
QY 243 IKANNLNKSDKRTKITTDKI-----EINLGFVQDTSTKRELEQIASKDKSTLAVAGPL 298
DB 175 AKYAN-----SNIKVVSQYVGTGDFGLG-----RSTASNN-----YRDGVDIIFPAAG-L 219
QY 299 TEIFSDIILANO--NDRYLIGVDIPOSIVYRTKTKKPFSTSLKNGYVSFVSLDLYKTKS 356
DB 220 SGIVIVAAKELGPDHIIIGVDQSYL---APRNIVYSAVKKYSLSLSTL-----TKKY 271
QY 357 NSRNLAGEFEGKKSATVYLGIK-DREVDIADTSLEGNDKRLATEAI 401

```

```
Db      272 LETGV--TDFGRK--TMFLGLKEDGIGLVNLETKNSYEIYKSL 312
RESULT      8
ID       TCSSA_LISMO          STANDARD:          PRT:          357 AA.
AC      048754;
DT      15-DEC-1998 (Rel. 37, Created)
DT      01-MAR-2002 (Rel. 41, Last sequence update)
DE      01-MAR-2002 (Rel. 41, Last annotation update)
DE      CD4+ T cell-stimulating antigen precursor.
GN      TCSSA OR LM001388.
OS      Listeria monocytogenes.
OC      Bacteria; Firmicutes; Bacillus/Clostridium group;
CX      Bacillus/Staphylococcus group; Listeria.
NX      NCBI_TaxId=1639;
[1]
SEQUENCE FROM N.A.
RN      STRAIN=ECD-e / Serovar 1/2a;
RX      MEDLINE=21537279; PubMed=11679669;
RA      Glaser P., Frangoul L., Buchrieser C., Rusniok C., Amend A.,
RA      Baqerlo F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA      Chaabot A., Chetouani F., Couve E., de Darvar A., Deloux P.,
RA      Domenech E., Dominguez-Bernal G., Duchaud E., Durant L., Dussauget O.,
RA      Entlan K.-D., Fsihl H., Garcia-del Portillo F., Garrido P.,
RA      Gautier L.-M., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA      Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kuypkatt G.,
RA      Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA      Nordstiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA      Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA      Vazquez-Boland J.-A., Voss H., Wehlund J., Cossart P.;
RA      "Comparative genomics of Listeria species.";
RL      Science 294:849-852(2001).
[2]
SEQUENCE OF 1-252 FROM N.A.
RP      STRAIN=85EO-1167;
RX      MEDLINE=96096448; PubMed=7500019;
RA      Sanderson S., Campbell D.J., Shasiri N.;
RA      "Identification of a CD4+ T cell-stimulating antigen of pathogenic
RA      bacteria by expression cloning.";
RL      J. Exp. Med. 182:1751-1757(1995).
-1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
(Possible).
-1- SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-slb.ch/announce/slb.ch)
CC      or send an email to license@isb-slb.ch).
CC
DR      EMBL: AL591979; CAC99466.1; -
DR      EMBL: S80336; BAB35725.2; ALT_TERM.
DR      ListList: LM001388; -
DR      InterPro: IPR003760; Bmp.
DR      Pfam: PF02608; Bmp; 1.
DR      PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
KW      Antigen; Membrane; Lipoprotein; Signal; Complete proteome.
FT      SIGNAL              1           22         PROBABLE.
FT      CHAIN               23           357        CD4+ T CELL-STIMULATING ANTIGEN.
FT      LIPID               23           23         N-ACYL DIGLYCERIDE (PROBABLE).
FT      SEQUENCE          357 AA; 38415 MW; 83605BB86419C8D1 CRC64;
SO
Query Match      7.4%; Score 169; DB 1; Length 357;
Best Local Similarity 23.8%; Pred. No. 0.0012;
Matches 98; Conservative 63; Mismatches 144; Indels 106; Gaps 22;
```

Db	19	ILGAGSSSDDKSSDDKSSKDF--TVAAVDTGVDSDRSFNQSAWEGLOKFGKANDMEK	76
Qy	110	GGELTSSVSSSTAELEGKYSSTLANTKKNWYLS-----GQHGDAFLRWLK-IPENQ	160
Db	77	GTGYNYVLOSASEADYK-----TYLNLNVAVRSDYDLVIGICIKLDALEEVSKORPKNP	130
Qy	161	LTETENIITILIGIDMTDENVIPTGRINLTYTEAGWLAGVANAISFLAKKFPSPDTKRS	220
Db	131	AIYDPT-----DDRNVVSGIG-----FNDNGSYLVGVVAGL-----TTKTN	168
Qy	221	AIIVGGISPAVTD-PIAGYLAGIRAMLNKSDKRTKITYTDKIEINLGFVDV--DTSK	276
Db	169	KGFVGKGVKGYTIDRFEGFAGTAGVKA-----VBNQAQIDVGYANDFAKA	212
Qy	277	ERLEDIASKDRS-----TLNAAVGLTEFSDITIANQDR-----YLIGVPTDQ-----	321
Db	213	DKGQIASMSVSGVDYTFHAAGVGNGVFAR-ARNLKKDKPSRAVWVYGVDRDQDECK	271
Qy	322	SLVYTKTKNKFFETSLKNLGYSVFSVLSDLTYTKNSRNLAG-PEFGKSAFVYLGID	379
Db	272	VTANGCKDYNTVLTSEIKRVDAV-----EDLVT-----RAKAGDPGCGK---IEYGIDK	319
Qy	380	RFVADIATSLLEGNDKKLATEAISLSEAKKEFEETKTIPAEVRYKTELEIPMP	430
Db	320	DAVGISE-----HODNISNDVLAKE---EYKQKIVDGD---IKVPERP	357
RESULT	9		
BMPA_BORAF		STANDARD: PRT: 339 AA.	
ID	BMPA_BORAF	O31280; O31281; O31282; O31283;	
AC	15-DEC-1998	(Rel. 37, Created)	
DT	15-DEC-1998	(Rel. 37, Last sequence update)	
DT	15-DEC-1998	(Rel. 37, Last annotation update)	
DE		Basic membrane protein A precursor (immunodominant antigen P39).	
GN	BMPA.		
OS	Borrelia afzelii.		
OC	Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.		
OX	NCBI_TaxID=29518;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=PKO, PMUDI, PLE, AND PLJ7;		
RC	MEDLINE=98010210; PubMed=9350727;		
RA	Roesler D., Hauser U., Wilske B.;		
RT	"Heterogeneity of bmpA (P39) among European isolates of Borrelia		
RT	burgdorferi sensu lato and influence of interspecies variability on		
RL	serodiagnosis.";		
RL	J. Clin. Microbiol. 35:2752-2758(1997).		
CC	-1- FUNCTION: NOT KNOWN; IMMUNOGENIC PROTEIN.		
CC	-1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor		
CC	(Probable).		
CC	-1- SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.		
CC			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to licensese@isb-sib.ch).		
CC			
DR	EMBL; X81516; CAA57236.1; -		
DR	EMBL; X97237; CAA65876.1; -		
DR	EMBL; X97239; CAA65878.1; -		
DR	EMBL; X97241; CAA65880.1; -		
DR	InterPro: IPR003760; Bmp.		
DR	Pfam: PF02608; Bmp. 1.		
DR	PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.		
KW	Antigen; Membrane; Lipoprotein; Signal.		
FT	SIGNAL	1	17
FT	CHAIN	18	339
FT	LIPID	18	18
FT	VARIANT	125	125
			BASIC MEMBRANE PROTEIN A.
			N-ACYL DIGLYCERIDE (PROBABLE).
			A -> S (IN STRAIN PLE).


```

Db 286 YGLKGVGV-----RNPKMISF-----ELKEIDNLSKTIKKEIIVP----- 325
Qy 435 DKQOESLDKLYND 447
Db 326 -SNKSEYEKFLKE 337

RESULT 11
Y040_MYCGE STANDARD: PRT; 591 AA.
ID Y040_MYCGE
AC P47286;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical lipoprotein MG040 precursor.
GN MG040.
OS Mycoplasma genitalium.
Bacteria: Firmicutes; Bacillus/Clostridium group; Mollicutes;
Mycoplasmataceae; Mycoplasma.
NCBI_TaxID=2097;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 33530 / G-37;
RX MEDLINE=96026346; PubMed=7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kiehlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.L.,
RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of Mycoplasma genitalium."
RL Science 270:397-403(1995).
[2]
RP SEQUENCE OF 448-517 FROM N.A.
RC STRAIN-ATCC 33530 / G-37;
RX MEDLINE=94075230; PubMed=8253680;
RA Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III;
RT "A survey of the Mycoplasma genitalium genome by using random
sequencing."
RL J. Bacteriol. 175:7918-7930(1993).
-1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
(CC (potential)).
CC (potential).
CC -1- SIMILARITY: SOME, TO T.PALLIDIUM TMPC.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

```

```

Db 33 ATKITDSSVQLLVSNFTLADKSFQMSYBGINSEFFKSKGVDDLEADSSQLQ-EGNGL 91
Qy 127 -----YSLANTFNKNVWVSGFQHGDAFTRWLKIPEKNOLFTEKNIILGI 172
Db 92 WKRPGETLSDRIATFNNNIKNGSDVIYATGFMQ-----QESLQAITSDIRF--- 138
Qy 173 DMTDTENVIPFG-----RYINLYKTEDEAGMLAGYANAS 206
Db 139 -OSDKESLAKTGFIFVDAIEKEFNKRNQVDPFKSTPNINISVAFRSDGSLFGVATAV 197
Qy 207 FL-----AKKPSDPTKSAIVIGG--GIS-PATYDFAIAGLAKIMN-----L 248
Db 198 YINLNOEYFLDKSGWSTNSNNNELTSGVGIAPLSLFLNGRLGIAFNIEYIYHNL 257
Qy 249 KNS-DKTKRTYTDK-----IEINLG-----FDVOD-----YST 275
Db 258 SDAQDSNAQVTTTSKQTVLKOQVANGEKRIKKIWIISPKQSDGRTINIQHQSGSFSDT 317
Qy 276 KERLEQIASK--DK-PSLLAVAGPFE-IFSDITANONDRYLICVDTQSL---VYTK 327
Db 318 EPRAITTIANNLIDKGVNAIIPACQPTNLVYQIARROAHTVAVIGVDSAOELDINIDAP 377
Qy 328 TRNKF-----FTSILKNIGVSPSVLDLYTKKNSNLAGFE-FGKKSATVYLG- 376
Db 378 NKDKLMGNKKKTIIPSSI-KALDVAVESILSTL--EKQSSON--GYQGFYNN---IGT 428
Qy 377 IKDRFVADIADT-----SLEGNDRKLTAEAISKAEKFEETK 413
Db 429 VKNNSVGVSSEAGYERFLIDPEVFNKNTSNQAMSLSLANNAASSDNNKKKISEVATK 485

RESULT 12
BMPB_BORAF STANDARD: PRT; 341 AA.
ID BMPB_BORAF
AC 031284;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Basic membrane protein B precursor.
GN BMPB.
OS Borrelia afzelii.
CC Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.
NCBI_TaxID=29518;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN-PKO;
RX MEDLINE=98010210; PubMed=9350727;
RA Roessler D., Hauser U., Malske B.;
RT "Heterogeneity of BmpB (P39) among European isolates of Borrelia
burgdorferi sensu lato and influence of interspecies variability on
serodiagnostics."
RL J. Clin. Microbiol. 35:2752-2758(1997).
-1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
(CC (probable)).
CC (probable).
CC -1- SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

```

Query Match 6.7%; Score 154; DB 1; Length 341;
 Best Local Similarity 25.4%; Pred. No. 0.0094;
 Matches 90; Conservative 56; Mismatches 125; Indels 84; Gaps 19;

87 GYVNDNSFNQSSWEALIOGLATGGEITSVDSTAELEKYS--LANTKN--VW 138
 36 GYLDDKSFNSANEALLRLKKDEPENIEKVFSSA--VSGYSSVSISLDLKNKSGSLW 93
 139 YLSPGHG-AFTRMKTIENKOLFTEKNII--ILGIDMTDENIVITGTIVMLTKTEE 195
 94 -LVGVLAVSISVSLNEP---INVGIDPIYGDVQIPYNL-----IGVERIRID 141
 196 AGVLGAYANASFLAKFPSPDPTKRSATVIGGISPAVTD-FINGYLAGIKAMN-----L 248
 142 GAFLAGY-----IAKKSYS-----SKIGFVGKGDIVAFRGYENGAKYANKGIEYS 192
 249 KNSDKTKITTDKIEIN---LGFVDVDTSTKERLEQIASKDPSTLLAVAGPLTEIFSD 304
 193 EYSNFSFSDIDIAVMANKYSKIGIDIHFPA--GLAGVGTEAPKEI----- 237
 305 ILANQDRLLIGVDQSLVYTKTKNKFSTILKNIGYSVSVLSLDTYKKSRLAGF 364
 238 ---GGYIVIGADQSHLAP---NFTTSVIAKNVGDALYLITSE-----SLKDNVW 284
 365 ERGKKSATVYLGIDRFVDIADTSLSEGNDKLATEAISEAKKEFEETKTIPAE 419
 285 EGGK---IYOMGLRDGVGLSNAN-----EFYIKDLERIKINKEIIVPCNQ 328

RESULT 13
 BMAPC_BORBU STANDARD: PRT: 353 AA.

AC 050169: 044859: (Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Basic membrane protein C precursor.
 GN BMAPC OR B80384.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 [1]
 RP SEQUENCE FROM N.A.
 RP STRAIN=297;
 RX MEDLINE=95080623; PubMed=7988902;
 RA Aron L., Aleksun M., Perlee L., Schwartz I., Godfrey H.P.,
 RA Cabello F.;
 RA "Cloning and DNA sequence analysis of bmapc, a gene encoding a
 RA potential membrane lipoprotein of Borrelia burgdorferi.";
 RA FEMS Microbiol. Lett. 123:75-82(1994).
 [2]
 RN SEQUENCE FROM N.A.
 RN STRAIN=297;
 RC MEDLINE=98065943; PubMed=9403685;
 RX Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
 RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
 RA Douherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
 RA Peterson J., Kierlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
 RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
 RA Ullrichback T., Matthey L., McDonald L., Artlich P., Bowman C.,
 RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
 RA Smith H.O., Venter J.C.;
 RA "Genomic sequence of a Lyme disease spirochete, Borrelia
 RA burgdorferi.";
 RL Nature 390:580-586(1997).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 CC (Probable).
 CC -1- SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U49938; AAC44711.1; -;
 CC EMBL: AF001143; AAC66756.1; -;
 DR TIGR: B80384; -;
 DR InterPro: IPR003760; Bmp.
 DR Pfam: PF02608; Bmp; 1.
 DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 KW Membrane; Lipoprotein; Signal; Complete proteome.
 FT SIGNAL 1 16 PROBABLE
 FT CHAIN 17 353 BASIC MEMBRANE PROTEIN C.
 FT LIPID 17 17 N-ACYL DIGLYCERIDE (PROBABLE).
 FT VARIANT 179 179 V -> L (IN STRAIN 297).
 SQ SEQUENCE 353 AA; 39824 MW; 2E8PF607D6CAB9B0 CRC64;

Query Match 6.4%; Score 147; DB 1; Length 353;
 Best Local Similarity 23.7%; Pred. No. 0.027;
 Matches 84; Conservative 53; Mismatches 131; Indels 86; Gaps 18;

83 GYVNDNSFNQSSWEALIOGL-----ALNGEITSVDSTAELEK 127
 34 VLAHSGYDYGKGYQSVHGVKLRDNFGIKLTKSLRPYIEKRLITDE--AMTEDAT 91
 128 SSLANTNKNVWVLSGPHGDAFTRWLKIPEKNOLFTEKNIIIGI---DWTDTENVITP 183
 92 EYQKNPLNFWL-LGYRFSDL-----SVLSTERPIYGIIDAFYGDIG--VPR 139
 184 GRINLTYTEEGWLAGVYANASFLAKFPSPDPTKSAIYIG--GGISPATVDFIAGYL 240
 140 NS-LAIKFRNEEAFLAGYIAAKMSRKE-----KIGFLTGMSRHWDFRFGFK 187
 241 AGIKAMN--LNSDKTKKITTDKIEINLGFVDVDTSTKERLEDA---SKDPSLLAV 294
 188 AGIFYANPKRLVSKRAPSLEF-----KEGKAMALFMKEDVGVYIFPI 232
 295 AGPLTEIFSDIIANQDR-VLIGVDVDTQSLVYTKTKNKFSTILKNIGYSVSVLSLDLYT 353
 233 AGITGLGVYAAKELGPKYVIGLNDQSYI---APQNVITSLIKIKGKIYISISEYT- 288
 354 KKSNSRLAGFEFGKKSATVYLGIDRFVDIA-DTSLSEGNDKLATEAISEAKK 406
 289 ---NNRVFKG-----GIIDRGLKRGVLEIVADPVLNN--RLVDEVIDLENK 331

RESULT 14
 PEBB_STRPY STANDARD: PRT: 498 AA.

AC 099XSL;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable dipeptidase B (EC 3.4.-.-).
 GN PEBB OR SP22066.
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1314;
 RN [1]
 RP SEQUENCE FROM N.A.
 RP STRAIN=SP370 / ATCC 700294 / Serotype M1;

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2002, 09:14:20 ; Search time 2:25 Seconds
(without alignments)
495.099 Million cell updates/sec

Title: US-09-676-249A-2
Perfect score: 2299
Sequence: 1 MKRRKRWKFLGLVLPPLS.....KQPDKQHSIDKLTIDINKN 451

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Minimum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	203	8.8	328	4	US-08-961-083-8
2	179	7.8	341	1	US-08-396-957A-5
3	162	7.0	339	1	US-08-396-957A-4
4	144	6.3	353	4	US-08-313-412-1
5	141.5	6.2	889	4	US-09-336-447A-15
6	131.5	5.7	1786	4	US-08-973-462-8
7	122.5	5.3	2285	4	US-09-308-375-2
8	122	5.3	2411	4	US-09-268-347-36
9	121.5	5.3	1104	4	US-08-923-992A-4
10	119.5	5.2	1861	2	US-08-790-912-4
11	119	5.2	892	4	US-09-336-447A-5
12	117.5	5.1	941	4	US-09-074-658-75
13	117	5.1	1912	1	US-08-409-995-4
14	117	5.1	1912	3	US-08-685-467-4
15	117	5.1	2353	4	US-09-377-155-33
16	117	5.1	2353	4	US-08-913-942-4
17	117	5.1	2353	4	US-09-669-974-33
18	117	5.1	2354	4	US-09-268-347-47
19	116.5	5.1	873	4	US-09-336-447A-13
20	115.5	5.0	1164	4	US-08-923-992A-2
21	115.5	4.9	466	4	US-08-235-836C-107
22	112	4.9	627	1	US-08-703-947-2
23	111.5	4.8	1098	1	US-08-409-995-2
24	111.5	4.8	1098	3	US-08-685-467-2
25	111.5	4.8	1098	4	US-09-377-155-32
26	111.5	4.8	1098	4	US-08-913-942-2
27	111.5	4.8	1098	4	US-08-923-992A-8

28	111.5	4.8	1098	4	US-09-669-974-32	Sequence 32, Appl
29	111.5	4.8	1098	4	US-09-268-347-44	Sequence 44, Appl
30	111	4.8	1128	4	US-08-923-992A-6	Sequence 6, Appl
31	110.5	4.8	1164	4	US-08-923-992A-10	Sequence 10, Appl
32	110.5	4.8	1964	2	US-08-790-912-3	Sequence 3, Appl
33	110.5	4.8	2052	2	US-08-790-912-2	Sequence 2, Appl
34	110	4.8	1579	3	US-08-755-587-184	Sequence 184, App
35	109.5	4.8	607	1	US-08-409-995-6	Sequence 6, Appl
36	109.5	4.8	607	3	US-08-685-467-6	Sequence 6, Appl
37	109.5	4.8	607	4	US-08-913-942-6	Sequence 6, Appl
38	108.5	4.7	693	4	US-08-235-836C-68	Sequence 68, Appl
39	108	4.7	941	4	US-09-336-447A-9	Sequence 9, Appl
40	107.5	4.7	693	4	US-08-235-836C-72	Sequence 72, Appl
41	107	4.7	703	3	US-08-646-273-25	Sequence 25, Appl
42	107	4.7	753	3	US-08-646-273-36	Sequence 36, Appl
43	106.5	4.6	894	2	US-08-867-941-15	Sequence 15, Appl
44	106.5	4.6	1183	2	US-08-447-031A-2	Sequence 2, Appl
45	106.5	4.6	2432	4	US-09-074-658-15	Sequence 15, Appl

ALIGNMENTS

```

RESULT 1
US-08-961-083-8
; Sequence 8, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 Inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; TELEPHONE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 328 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-961-083-8

```

Query Match 8.8%; Score 203; DB 4; Length 328;
Best Local Similarity 25.7%; Pred. No. 1.5e-10;
Matches 98; Conservative 56; Mismatches 144; Indels 84; Gaps 19;

QY 61 RKSEIMAKADANKFGLNMAIYAGTVDNNSWSEAIQOLGA---LTGGEITSVD 117

```

Db 3 RSSRNAASSDVK-----TKAIVTDTGVDKSFNOSANEGLQANGKEHNLKNGCTYF 58
Oy 118 SSTAELE-----GKYSSLANTNKNWVLSGFQHGDAFTRLKIPENKQLETEKNII 168
Db 59 OSTSEADYANNLQQAAGSYNLIFGV-----GFLANNA-----VKAAAEHNDLVANV 104
Oy 169 ILGIDMTPEENIIPGRYINLTYTEEAGWLAGYANASFLAKFPSPDKRAIYIGCI 228
Db 105 LIDVYIKDQKN-----ASVTFADNESGYLAGVAAK-----TKTKQVGFVGCI 149
Oy 229 -SPAVTDFIAGYLAGIKAWNLKNSDKTKITDKIEINLGFVDQDSTKERLEQJASKDK 287
Db 150 ESEVTSRPFAGKACV-----ASVDSIVQVDYAG-SFG-DAAGKTIAMAOYAGAD- 201
Oy 288 PSTLLAVAGPL-TEIFSDIIANQND-----YLIGVDTQSL-VYT-----KTKNKFT 334
Db 202 -IYVQVAGGTGAGVFAEAKSLNESRPENEKYVIGVDRDQEAEGKTYTSKDKESFVLV 259
Db 335 SLKMLGYSVFSVSLDLYTKKSNRNLAGFEFGKKSATVYLGIKDRFVDIADTSLGNDK 394
Db 260 STLKQVGTIVKDI-----SNKAERGFPGQ---VIVSLKDKGYDLAVTNLSEEGK 308
Oy 395 KLAIEAISEAKKEFEKTKTIP 416
Db 309 K-----AVEDAKKILDGSKVP 326

```

RESULT 2

US-08-396-957A-5

Sequence 5, Application US/08396957A

Patent No. 5780041

GENERAL INFORMATION:

APPLICANT: SIMPSON, WARREN; SCHWAN, TOM G.

TITLE OF INVENTION: ANTIGENIC PROTEINS AND

TITLE OF INVENTION: GENES ENCODING SAME OF BORRELLIA BURGDORFERI.

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN & FINNEGAN, L.L.P.

STREET: 345 PARK AVENUE

CITY: NEW YORK

STATE: NEW YORK

COUNTRY: USA

ZIP: 10154

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WORDPERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/396,957A

FILING DATE: 01-MAR-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/664,731

FILING DATE: 05-MAY-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/487,716

FILING DATE: 05-MAY-1990

ATTORNEY/AGENT INFORMATION:

NAME: DOROTHY R. AUTH

REGISTRATION NUMBER: 36,434

REFERENCE/DOCKET NUMBER: 2026-4018054

TELEPHONE: (212) 758-4800

TELEFAX: (212) 751-6849

TELEPHONE: (212) 751-6849

TELEFAX: (212) 751-6849

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 341

```

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: No
ORIGINAL SOURCE:
ORGANISM: Borrelia burgdorferi
STRAIN: Sh-2-82
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
FEATURE:
NAME/KEY: p39a
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION: p39a protein sequence
US-08-396-957A-5

```

Query Match 7.8%; Score 179; DB 1; Length 341;
 Best Local Similarity 26.9%; Pred. No. 2.6e-08;
 Matches 101; Conservative 59; Mismatches 123; Indels 92; Gaps 25;

```

Oy 72 ANKHFGILMAIYTAGTIVNDNSFNOSMEAIQGLALGTGETSYDSSSTAELEKYS-- 129
Db 24 SSKKIKISMV---DGVLDKSFNSANBALRLKKDPENIEEYFS--CAISGVYSSIV 78
Oy 130 --LANTNKN---VWVLSGFQHGDAFTRLKIPENKQLETEKNII--ILGIDMTPTENVI 181
Db 79 SPLDMLKRRGSDLI--LVGYMLTD--SLVASENPKI--SVGIIDPIYGDVOJPEKLI 133
Oy 182 PGRYINLTYTEEAGWLAGY--ANASFLAKFPSPDKRAIYIG--GGISPAVTD-FI 236
Db 134 A-----VFEREAGAFIAGYIAAKSFSGK-----IGFGCMGMGNIVDAFR 174
Oy 237 AGYLAKAMNLKNSDKTKITTDKIEINLGFVDQDSTKERLEQJASK--DKPSTLLAV 294
Db 175 YCEBGAAYAN-KDIEITSEVNSFSVDIG-----RTYASKMWSKGIQVYHF 221
Oy 295 AGPLTEI-FSDIIANQND-RYLIGVDTQDQSLVYTKRKPFSLKNLGYSVFSVSLDLY 352
Db 222 AAGLAGIVTEAKNMGIGYVIGADQDS--YLAPKN-FITSYVKNIGDALYLITGE-Y 277
Oy 353 TKKSNRNLAGFEFGKKSATVYLGIKDRFVDIADTSLGNDKKLATEAISEAKKEFEK 412
Db 278 IKNNV-----WEGK---VVQMGIRDVIGLPPAN-----EF-EYI 310
Oy 413 KTIPEEYKTLPI 427
Db 311 KVLKRIYVKEITVP 325

```

RESULT 3

US-08-396-957A-4

Sequence 4, Application US/08396957A

Patent No. 5780041

GENERAL INFORMATION:

APPLICANT: SIMPSON, WARREN; SCHWAN, TOM G.

TITLE OF INVENTION: ANTIGENIC PROTEINS AND

TITLE OF INVENTION: GENES ENCODING SAME OF BORRELLIA BURGDORFERI.

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN & FINNEGAN, L.L.P.

STREET: 345 PARK AVENUE

CITY: NEW YORK

STATE: NEW YORK

COUNTRY: USA

ZIP: 10154

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK
 COMPUTER: IBM PC COMPATIBLE
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WORDPERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/396,957A
 FILING DATE: 01-MAR-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/020,245
 FILING DATE: 19-FEB-1993
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/664,731
 FILING DATE: 05-MAY-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/487,716
 FILING DATE: 05-MAY-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: DOROTHY R. AUTH
 REGISTRATION NUMBER: 36,434
 REFERENCE/DOCKET NUMBER: 2026-4018054
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 758-4800
 TELEFAX: (212) 751-6849
 TELEX: 421792
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 339
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: Borrelia burgdorferi
 STRAIN: Sb-2-82
 INDIVIDUAL ISOLATE:
 DEVELOPMENTAL STAGE:
 HAPLOTYPE:
 TISSUE TYPE:
 CELL TYPE:
 CELL LINE:
 ORGANELLER:
 FEATURE:
 NAME/KEY: p39
 LOCATION:
 IDENTIFICATION METHOD:
 OTHER INFORMATION: p39, protein
 OTHER INFORMATION: sequence.

[illegible]

```

0Y  315  IGVDDIDSLVYTKRKNNKFEETILNKNGSVSVSLDLYTKKSNRNLAGPEFGKSAITY 374
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  239  IGVDDQAYL---APDNVYITTDVGRAL-----NFT---SNLTKNTFEGG--LIN 285
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
0Y  375  LGIKRPFADIDATSLGNDKRLKATEAISAKKEFEETKTTPAEVRRKTEIPMPDKP 434
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  286  YLGEKVGVGV---RNPKMTS-----ELEKELNLSRKINKETIIV----- 325
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
0Y  435  DKODESLDKLITD 447
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  326  -SNKSEYKEFKLE 337
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

```

RESULT 4
US-08-313-412-1
: Sequence 1, Application US/08313412
: Patent No. 6248583
: GENERAL INFORMATION:
: APPLICANT: Aron Ph.D., Lieselotte
: APPLICANT: Cabello M.D., Felipe
: APPLICANT: Godfrey M.D., Henry P.
: APPLICANT: Schwartz Ph.D., Ira
: TITLE OF INVENTION: Chromosomally-Encoded Membrane
: TITLE OF INVENTION: Protein of Borrella burgdorferi
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Nixon, Hargrave, Devans & Doyle
: STREET: P. O. Box 1051, Clinton Square
: CITY: Rochester
: STATE: New York
: COUNTRY: USA
: ZIP: 14603
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/313,412
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Goldman, Michael L.
: REGISTRATION NUMBER: 30,727
: REFERENCE/DOCKET NUMBER: 35553/1020
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 716-263-1304
: TELEFAX: 716-263-1600
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 353 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-313-412-1

```

Query Match	6.3%	Score 144	DB 4	Length 353
Best Local Similarity	23.4%	Pred. No. 4.4e-05		
Matches	83	Conservative	54	Mismatches 131; Indels 86; Gaps 18;
OY	83	VTAGGTVDNSNSQSWEAIOOLG	-----ALNGEITVSSTAELECKY	127
Db	34	VLAHSSFDKGVQSVYHDSVAKLRNFQIKLTKLRPIEGKRLTYDE	--AMTEDMY	91
OY	128	SLANTNNKVVVLSGFQHDADFTRWLIKIPENKOLFTEKNIILIG	---DWTDTENVIPT	183
Db	92	EVOQNPLNLFWLI-GYRFSDL	-----SVKLSERPDIIYGIIDADYDGIQ--VPK	139
OY	164	GYVILYTKTEAGVLAGANNSFLAKKRPSPDKRSALVIG	---GGISPAVTDPLAGL	240
Db	140	NS-LAIKRNNEAALAGTIAAKMSRKE	-----KIGFLTPGMSHKLKDRFGKR	187

```
0Y      241 AGICKAWN--LKSUSDKKTKITTTQKEINLEIGEDVODTSTKRLEBIA-----SKDKPSUILLAV   294
        || | | :| :||
Db      188 ACIFANPRLRSLYSKRAPLFD-----KEKGRAMLFMTKREDRGVYIFFI    232
        || | | :| :||
0Y      295 ACPUTEIFSDIIANONDR-VLIGVDPTDOSLVYTKRKNEFFTSILKNLGVSVPISLDLYT   353
        || | | :| :||
Db      233 ACITTCGCVDAKAKEJGPCKYYVIGLNDOOSYI---APQNTYSTIKNKIDGVIVSISEYI-   288
        || | | :| :||
0Y      354 KXSNSRNLAGFEFGKKSATVVYGIRKPFDIA-DTSLSEGNDRKLATEAISEAKK   406
        || | | :| :||
Db      289 ---NNRVENG-----GIINDRGELGEVIETVKDPVLANN--RLNDEVIDLENK    331

RESULT          5
US-09-336-447A-15
; Sequence 15, Application US/09336447A
; Patent No. 6310190

GENERAL INFORMATION:
APPLICANT: HANSEN, ERIC J.
APPLICANT: AEBI, CHRISTOPH
APPLICANT: COPE, LESLIE D.
APPLICANT: MACIVER, ISOBEL
APPLICANT: FISKE, MICHAEL J.
APPLICANT: FREDENBURG, ROSS A.
TITLE OF INVENTION: URSAL AND USP#2 ANTIGENS OF MORAXELLA CATARRHALIS
FILE REFERENCE: AMCY:024
CURRENT APPLICATION NUMBER: US/09/336,447A
CURRENT FILING DATE: 1999-06-21
NUMBER OF SEQ ID NOS: 98
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 15
LENGTH: 889
TYPE: PRT
ORGANISM: Moraxella catarrhalis
US-09-336-447A-15
```

[illegible]

```

; RESULT 6
; US-08-973-462-8
; Sequence 8, Application US/08973462B
; Patent No. 6191270
; GENERAL INFORMATION:
; APPLICANT: DROUILHE, PIERRE
; APPLICANT: DAUBERGES, PIERRE
; TITLE OF INVENTION: MLARAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 0660-0125-0 PCT
; CURRENT APPLICATION NUMBER: US/08/973,462B
; CURRENT FILING DATE: 1998-02-06
; EARLIER APPLICATION NUMBER: PCT/FR96/00894
; EARLIER FILING DATE: 1996-06-12
; EARLIER APPLICATION NUMBER: FR 95/07007
; EARLIER FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1786
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Polypeptide
; US-08-973-462-8

```

```

Query Match Similarity      5.7% ; Score 131.5 ; DB 4 ; Length 1786 ;
Best Local Similarity      19.4% ; Pred No. 0.0084 ;
Matches      85 ; Conservative      75 ; Mismatches 145 ; Indels 133 ; Gaps 19

OY      37 EEKSDANONKQITDYSKISGLVNERKSEI-----MAAKDANKHGLNMAIYTAGG 87
          |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB      908 EKTQSEEEKKEIVD-----IEEVEEYATVLIETVEQAEKESAN-----TITEIFE 954
OY      88 TVDNMSF--NOSMWEIQOGLALTGGEITSVDSSTAELGKXYSILANTKKNVWVLSPQH 145
          |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB      955 NLEENAVESNEVAAENLEKLTNETFNTYVLDKYEEVEISGE--SLENNM-----1002
OY      146 GDAFRWLKIPENKQILTEKNIIILIGIDMTDENYIPTGRY-----INLYKTEEAQWLAG 201
          |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB      1003 -----DKAFESLFDNVAGIQ---ENLL-TGMPFSIERSIYIQSDE-----1039
OY      202 YANASFIAKKFSDPTKSAIYIGGIGISAVYDFIAGIAGI--KAMNIAKNSDKKTIIT 259
DB      1040 -----KVDLENVYSSILDNTENMKEGILNKLENISSTEGVETVY 1080
OY      260 DKIEINLFQVDYDTSTKERLEQIAKDKRSTLLAAGPTEIF---SDITIANONDRXYLAG 316
          |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB      1061 EHYEQENVYDV-----DYPAKKDQDLGLINELGSGIKEMFNFNEDYFKSSD--VIT 1129
OY      317 VDTDOSLYYTKTKNFFTSILKNLGSVFSVLSLDLYTKKSNRNLAGFEFGKKSATVYLG 376
          |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB      1130 VEEIDEPEQKVEKEVETYSIIIEEMENIYDVLEE--EKED-----1167
OY      377 IKDRVDYADTSLF-GNKKLATAEISEAKKE-----FEKKRTIPEAYRKTLEIPEMPD 431
          |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB      1168 LDKRIDVAEESITSSDSKETESTIKDEKDVSLVEEYQDNMDESVEKYELKNN--1225
OY      432 KQPDKQOESLKLINDIN 449
          |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB      1226 -----EEELMKDAVEIN 1237

RESULT      7
US-09-308-375-2
; Sequence 2, Application US/09308375
; Patent No. 6300117
; GENERAL INFORMATION:
; APPLICANT: Genencor International, Inc.
; TITLE OF INVENTION: Proteases from Gram-positive Organisms
; FILE REFERENCE: GC394-PCT

```

```

; CURRENT APPLICATION NUMBER: US/09/308,375
; CURRENT FILING DATE: 1999-05-14
; EARLIER APPLICATION NUMBER: EP9719636.4
; EARLIER FILING DATE: 1997-09-15
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 2285
; TYPE: PRF
; ORGANISM: Bacillus subtilis
US-09-308-375-2

```

```

Query Match          5.3%; Score 122.5; DB 4; Length 2285;
Best Local Similarity 19.3%; Pred. No. 0.083;
Matches 105; Conservative 84; Mismatches 188; Indels 167; Gaps 23;

```

```

OY 15 LVFPLSAITISAGCWDKETTKEKSADNONKQITDVSKISGLVNERKESIMAAKADANK 74
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 452 LNNINANDSISIA-----DKLINEVDNNNAVTTLDANSIRKAGSTASTFGEVLND 502
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 75 HEGLNMAIYTA---GGTVNDNSF-----NOSWEATOOLG---ALTGGETSVDS 119
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 503 LIGYTTAISTRESGNIVGNSLKTTFARIGNOSSIKALEQIGISYKTAGGAKSASDL 562
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 120 TAELEKYSLSANTNK-----NWVLSGF-----OHGDAFTFW 152
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 563 ISEVAKKMPILSDAQONMSIGVAGIYQLSRFAMNNFSAQMAKTAANSYGSASEQ 622
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 153 LKIPENKOL-----FTEKNIIIGIDMTDENVI---PTGRYINL-TYTEAGW- 198
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 623 OKYADSIQARVNKRONNTE--FAIASDAFISDGLIEFTQAAGSLMNSTGVYKISVGL 680
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 199 ---LAGYANASFLAKKPPSPPTKRSALVIG-----G3ISPA----- 231
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 681 PPLLAIVSTRITLISK--NRITLASSILIGTRAMGOETLTAAGLEAMTPAAVASRYLKT 738
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 232 -----VTDEFIAGYLAGIKAMNLK---NSDKKTKITTDKLEINLIGFVODSTKERLEQI 282
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 739 ALKGLIVSTLVGAFAPAL-GMALESILISSFAEAKKAKADPEQSOQRNVEAITTNK----- 792
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 283 ASKDKSTLLAVAGPLEIFSDIIANONRYLLGVPTDQSLVYTKKANKFTSILKNLGY 342
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 793 DSTDT-----LIQOKRELQKYESRSLTSDEOEYLYOTQOLAQTFPALVNGYDS 842
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 343 SVFSVLSDIYTKKSNRNLAGFEFGKKSATVYIGIKRPFDIADTSLGENDKKLATPAIS 402
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 843 QGNAIL-----KTNKELEKAIENITKE---YIALKKQ-----ETRD 874
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 403 EAKKEFEKTKTI--PAEEVAKTLEIPEMPDK-----OPDKOQESLDKL 444
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 875 SAKKTEDASKKEIKKSGDELKQYKQIADYNDKGRPKWDLIADDDYKVVAADKAKQSMILKA 934
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 445 ITDI 448
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 935 QSDI 938
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 8
US-09-268-347-36
; Sequence 36, Application US/09268347
; Patent No. 6335182
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS
; FILE REFERENCE: 1038-860
; CURRENT APPLICATION NUMBER: US/09/268,347
; CURRENT FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 36
; LENGTH: 2411
; TYPE: PRF

```

```

; ORGANISM: Haemophilus influenzae
US-09-268-347-36

```

```

Query Match          5.3%; Score 122; DB 4; Length 2411;
Best Local Similarity 20.8%; Pred. No. 0.1;
Matches 108; Conservative 65; Mismatches 220; Indels 126; Gaps 23;

```

```

OY 12 GLGLVPLSAITISAGCWDKETTKEKSADNONKQI-----TDSKISG-----LVN 59
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 303 GNGLVTRAKAIVDAVNKRAKGVKTT-----TANGONGDEPATYASGTNVTFFESGDGTASVTK 358
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 60 ERKSEIIMAAKADANKHEGLNM-----AIVTAGTVND----- 91
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 359 DTGNGCITVAVDAQVKGGLKFPDSDKIVADPTFALTVGVAAIEAKEDDKKLVLNAGDLV 418
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 92 NSRNOSWEALIQOLGALTGETSVDSSTAELEKYSLSLANTKNVWL---SGPOHG-- 146
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 419 TALGNLSWKAKAADTDTGDALEGI-SKQDEVKAGETVTEKAKKNLKVQDGANFTYSIQ 477
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 147 DAFTRMUKIP-----ENKOLFTEKNIIIL-----GIDMTDENVIPTGRYINLTYK 192
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 478 DALTGTLSTLGGTTNGNDAKTYINKDGLTTPAGNGGTTGNTISYTKDGIKAGKAI 537
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 193 TEAGWLAGYANASFLAKKPPSPPTKRSALVIGG3ISPAVTFEYAGYLAGIKAMNLKNSD 252
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 538 TNVASGLRAVDANF-----DVLNNSATDILNRHVEDA-----YKGLMLNEMKAN 582
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 253 KTKTKITTDKIEI-----NLGFDVO-DTSTKERLEQIASKRKPSLLAVAGPLEIFISDI 306
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 583 KQPLVDTSTAATVGDLEKLMVAVSTKNGTKEESQVQKQADE--VLTFGAQAAVATSK-- 637
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 307 ANONDRYLI-----GVPTDQSLVYTKKANKFTSILK--NLGYSVFSVLSDLTY 353
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 638 -SENGKTTIVSAETKADSGLEKDGDTIKLKVDNQTNDVLYVGNNGTAV-----T 688
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 354 KKSNSRNLAGFEFGKKSATVYLGIKDR-EVDIADTSLGENDKKLATPAISAKKEFEKT 412
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 689 K-----GGFEYVTKGAT---DADRGKVTYDATANDAKKVAAT--VADVATAINSAA 735
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 413 KTIPEAEVAKTLEIPEMPDKOPDKQDESILDKLITDINKN 451
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 736 TFWKTENLTTSIDEDNPTDNGKDALKAGDTLTFERAKGN 774
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 9
US-08-923-992A-4
; Sequence 4, Application US/08923992A
; Patent No. 6280738
; GENERAL INFORMATION:
; APPLICANT: Tai, Joseph Y.
; TITLE OF INVENTION: NO. 6280738-IGA Fc Binding Forms of the Group B
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/923,992A
; FILING DATE: 05-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,707

```

FILING DATE: 06-SEP-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Emond, Robert W.
 REGISTRATION NUMBER: 32,893
 REFERENCE/DOCKET NUMBER: 1438.0140001/RWE
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 371-2600
 TELEFAX: (202) 371-2540
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1104 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-923-992A-4

Query Match 5.3%; Score 121.5; DB 4; length 1104;
 Best Local Similarity 18.2%; Pred. No. 0.032;
 Matches 86; Conservative 77; Mismatches 207; Indels 103; Gaps 13;

38 EKSADNONQITDVSKISGLVNERKSEIMAKADANKHFGGLMAIVTAGTVNDNSFNOS 97
 164 DEELNHQKQVEMAMQAITNEDKSMKTIEDIRKO-----AQAADKKEDA 211
 98 SWEAIDQALGTGETISVDSSTAE-LEGKYSLANTNKNVWVLSGFQHGDAFTRMKIP 156
 212 EVKVRRELGLKLFSSFKAGLDQOIOEHVKKETSSEENTQ-----KVD 252
 157 E-----NKQFTEKNITILIDTDTENVIPTRGYINLYTKTEBAGVLAYANASFLAKF 212
 253 EHYANSLQALQKSLLEELKATTEGATGVKNOFLENAOKLEIOLIKETVKKLY--KA 310
 213 PSDPTKRSALVIGGISPATVDFIAGYLAGIKAMNLKNSDKTKITTDKIEINLGFVD 272
 311 MESLEQVKEKELHNSANLQDLVAKSEIYREBGLKNSKLPKLE-----EENH 365
 273 TSTKERLEQIASDKRSTLLAVAGPLTEIFSDIIANONDRYLI--GVDTQDSLVTYTKTN 330
 366 SKLKQVVEHFRKKFKTSEQVT---PKRYKRDLAANENNQOKTEILVSPENITVYGEDEV 422
 331 KF-----FTSLKMLGYSVFVSLDLYTKKNSRNLAGFEFG-----KSAIV 373
 423 KTTVAKSDSKTTLDFSLTLTKYNSVSDRISTNTKNTDNNKIAETITIKNLKLNQSYV 482
 374 YGIDRFVADIADTSL-----EGNDKKLATEAIS 402
 483 TIKANDDSGNVEKFTTIVQKKEQVKTPEQKHSKTEQVNPQEPKSNKNOLOELIK 542
 403 EAKKEFEKTKTTPAEVVRKLTLEIPMPKOPK-----QOESLDKLITDINK 450
 543 SAQOELKLEKAI--KELMEQPEIPSNPEYGIQKISWESQKEPIQOALISFVK 593

RESULT 10

US-08-790-912-4

; Sequence 4, Application US/08790912

; Patent No. 5976542

; GENERAL INFORMATION:

; APPLICANT: Weiser, Jeffrey N.

; APPLICANT: Plaut, Andrew G.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESS: PANITCH SCHWARZE JACOBS & NADEL, P.C.

; STREET: 1601 Market Street, 36th Floor

; CITY: Philadelphia

; STATE: Pennsylvania

; COUNTRY: USA

; ZIP: 19103-2398

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/790,912
 FILING DATE: 29-JAN-1997
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/026,939
 FILING DATE: 23-SEP-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Leary, Kathleen
 REGISTRATION NUMBER: 36,317
 REFERENCE/DOCKET NUMBER: 7600-401
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (215) 567-2020
 TELEFAX: (215) 567-2991
 TELEX: 831-494
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1861 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-790-912-4

Query Match 5.2%; Score 119.5; DB 2; length 1861;
 Best Local Similarity 18.1%; Pred. No. 0.11;
 Matches 95; Conservative 63; Mismatches 172; Indels 195; Gaps 21;

16 VFPLSAIATISACMDKETEKEESADNONQITDVSKISGLVNERKSEIMAKADANKH 75
 1355 VEVISTISLGGADRYRNNNEHRAGAELNK-----FVEDNAQETARQND--H 1401
 76 EGLNMAIVTAGTVNDNSFNOSWEAIDQALGTGETISVDSSTAELEGKYSLANTNK 135
 1402 YVYWRIL-----DEQREKLYRNILVYDAYKFGDDTVDKAVVDAQ----- 1443
 136 NWVLSGFQHGDAFTRMK-----KIPENKO--LTFEKIIILGIDTDTENVIPTRGYI 187
 1444 -----FDSSNPAMKYEFGPGVGNKVVHNRKGAVATGDSVYMGYRMLDKDAITYTH-- 1494
 188 NLTYKTEPAGWLAGYANASFLAKFSPDPTKRSATVIGGISPATVDFIAGYLAGIKAMN 247
 1495 EMTHSDNFIYLGIGRGRGLGPER-----FAKGLL-----Q 1526
 248 LNSDKTKITTDKIEINLGFVDITSTKERLEQIASDKRSTLLAVAGPLTEIFSDIYA 307
 1527 ABDHPDADATITVNSI---LKYDKNDASEKSRLOVL---DPTKRQNDL---KNYVH 1575
 308 NOND-----RYLIGV-----DTQDSLVTYTKKKEF-----TSILKLN----- 340
 1576 NMEVDIYMLEYLEGMSIVNRLSDVQVNALRKIEKNKYVADADGNDVYATNVIKNTMDA 1635
 341 -----GYSEFVSLDLYTKKNSRNLAGFEFGKKSAT 372
 1636 OKLNSFNLSLENDILSAREYKNGDYERNCYHTIKLFSPIYLSLSEKGPGLMKRRILAY 1695
 373 VYL---GIDRFVADIADT-----SLEGNKKLATEAI-----SE 403
 1696 ELLAAKGFQDGVVPIYNSQYEDDAKQNGKTIISYKTRGLVTDVLKRVFNGCFNNMTE 1755
 404 AKKE-FEEK-----TKTTPAEVVKTL 424
 1756 FRKMYEERKKNFDSLNVKTFDDTRQPMWTSYATKTIISTVEELQTL 1800

RESULT 11

US-09-336-447A-5

; Sequence 5, Application US/09336447A

; Patent No. 6310190

```

; GENERAL INFORMATION:
; APPLICANT: HANSEN, ERIC J.
; APPLICANT: AEBI, CHRISTOPH
; APPLICANT: COPE, LESLIE D.
; APPLICANT: MACIVER, ISOBEL
; APPLICANT: FISKE, MICHAEL J.
; APPLICANT: FREDENBURG, ROSS A.
; TITLE OF INVENTION: USPA1 AND USPA2 ANTIGENS OF MORAXELLA CATARRHALIS
; FILE REFERENCE: AMCY:024
; CURRENT APPLICATION NUMBER: US/09/336,447A
; CURRENT FILING DATE: 1999-06-21
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 892
; TYPE: PR
; ORGANISM: Moraxella catarrhalis
US-09-336-447A-5

```

Query Match

Best Local Similarity 5.2%; Score 119; DB 4; Length 892;

Matches 105; Conservative 74; Mismatches 196; Indels 130; Gaps 23;

```

QY 23 AITSAQCMQETTKKESADNOKQIT-DVSKISGLVNERKSEIMAKKADANKHFGIMMA 81
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 146 STIGGGYVNOATGEKSTVAGRRNNOATGNNSTVAG-----GSYNQATGNNST 192
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 82 IYTAGGTVDNSFNQSSWEAIQOLGALTGEIITSSTAELGKYSLANTKKNVWLS 141
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 193 V--AGG-----SHNQATGE-----GSPAAGVENKANNANNAVALGKNNTIDGNS----- 234
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 142 GFOHGDFTRMWKIPENKQLETER-NIIIGIDMTDEN-----VIP 182
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 235 -----VAIGSNNTIDSKQNVFIIIGSS-TMTTNGQSSVLLGHTAKKATVAVS 282
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 183 TGRYINLTYKTEAGWLAGYANASFLAK--KFPSPDKSAIVIG-GGISPAVTDPIAG 238
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 283 SAKVNGILT-----LGNFAGASKTGNGTVSVSENNENQIYVNGAGNISADSTDVNG 334
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 239 -----YLAGIKAWNLKN-----SDKTKTKITDK--TEINL--- 266
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 335 SOLYALATAVKADADENFRALTYTQNTLLIOGAQDALIAQONOTDIYANTAIERNFNRT 394
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 267 ---GPDVQDTS---TKERLE-QIASKDKPSTLLAVAGPLTEIFSDITANONDRYLIGVDT 319
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 395 VVNGFEIEKKKAGIAKNQADIQITLNNVGBELLNLGRLDQKADIDNNINNIYDLAQOQ 454
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 320 DQSLVYTKRKNKFTSLKRLGYSVF---SVLSDLTYTKSNSR---NLAGEFPGKKS- 370
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 455 DQSSSDIKTLKKNVEBGLDLSGRLLIDQKADLRDKITLNNVEEGGLDLSGRLLDQKAD 514
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 371 -ATVYIGIKDRFYDIAD---TSLEGNDKRLATEAISEAKEFEKTKTTPAPE--VRKT 423
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 515 IAKNQADIAQONOTDIQDLAAYNELQOYAKQOTEAIDALNKKASSANTDRIATJELGIAEN 574
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 424 LEIPEMPDKQPKQOESLKLITDI 448
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 575 KQDAQIAKAQANENKQGIKKNQADI 599
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 12
; US-09-074-658-75
; Sequence 75, Application US/09074658
; Patent No. 6184371
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M
; APPLICANT: Run-Pan Du
; APPLICANT: Quljun Wang
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H
; TITLE OF INVENTION: LACROFERRIN RECEPTOR GENES OF MORAXELLA
; NUMBER OF SEQUENCES: 78

```

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/074,658
; FILING DATE: 08-MAY-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-795
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1153
; INFORMATION FOR SEQ ID NO: 75:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 941 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-074-658-75

```

Query Match

Best Local Similarity 5.1%; Score 117.5; DB 4; Length 941;

Matches 79; Conservative 68; Mismatches 157; Indels 101; Gaps 17;

```

QY 32 KETTKERKSDNOKQITVDKISG-----LVNERK-----SEIMAKAD 71
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 256 KQKKAEYFLAEGSEELKPAKLAGNNGNYLNQLRNRYVERKKNQSLSAEEEMVREAO 315
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 72 ANKHFGILMAIYVAGGVNONSFNQSSWEAIQOLGALTGE--ITSVDSSTAELGKYSS 129
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 316 A-RHENLSAQAATGGGILDPMDYRSGSWLAKIGYRFGGRHYVGGVFEPTQ--RYDI 371
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 130 LANTKKNVWL-----GFOHGDFTRMWK-IPENKQLETKNII-LGIDMTDE 178
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 372 RDMTEKQYITDEATKTRKSDSGVYDGDPRDGLYFVNIEMKQDKNLVKGIGLKYSRTK 431
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 179 NVIP--TGRYINLTYKTEAGWLAGYANASFLAKKFPSPDKSAIVIGGISPAVTDPI 236
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 432 FIDEHRRRRRGGILYRENEYSPDNWADKAVLSFDKQGVATDNTLTLCNCAVYPSVDK-- 489
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 237 AGYLAGIKAWNLKNSDK-----KTKITDKIEMLGPDVQDTSIKER 278
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 490 ACRASADKPYSDSDFFHYRBEQHNVLNLFERSLKKRKTWHHTLTGFGYD----- 540
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 279 LEOIASK--DKPSTLLAVAGPLTEIFSDITANONDRYLIGVDTQSLVYTKRKNKFTSI 336
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 541 -----ASKAAYSREQLSHMARISE-FSDYADDGKRYKLLK-----PEV 579
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 337 LKNLGYSVFSVLDTYTKKSNRLAGFPGKKSATVYLGIKORF 381
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 580 VEG---SVCGYIETLRSRKCVPRKING-----SNHISLNDRF 614
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 13
; US-08-409-995-4
; Sequence 4, Application US/08409995
; Patent No. 5646259
; GENERAL INFORMATION:
; APPLICANT: Barenkamp, Stephen I.
; APPLICANT: St. Gene III, Joseph W.
; TITLE OF INVENTION: Haemophilus Adhesion Proteins

```

```

: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
: STREET: Four Embarcadero Center, Suite 3400
: CITY: San Francisco
: STATE: CA
: COUNTRY: USA
: ZIP: 94111-4187
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/409,995
: FILING DATE: 24-MAR-1995
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Silva, Robin M.
: REGISTRATION NUMBER: 38,304
: REFERENCE/DOCKET NUMBER: A-61053/RFT
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 781-1989
: TELEFAX: (415) 398-3249
: TELEX: 910 277299
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1912 amino acids
: TYPE: amino acid
: STRANDEDNESS: double
: TOPOLOGY: unknown
: US-08-409-995-4

```

Query Match 5.1%; Score 117; DB 1; Length 1912;

Best Local Similarity 21.0%; Pred. No. 0.2; Indels 128; Gaps 24;

Matches 109; Conservative 65; Mismatches 217;

```

OY 12 GIGLVFPLSATATTISAGCWDKETEKEKSADNQNQI-----TDVSKISG-----LVN 59
DB 306 GNGLTATAKAVIDAIVKAKGRVKT-----TANGQNGDFATVAGSTVTEPSGGTTASVTK 361
OY 60 EKKSSEIMAKADANKHFGILN-----AIVTAGTVND----- 91
DB 362 DTNGGITYVKKADAKGDDGKFPDSDKKIYADTTALTVTGGKVAEIAKEDDKKLVNAGDLY 421
OY 92 NSFNOSSWEAIDQALNGEITTSVDSSTAELGKYSLSLANTNKNVWYL---SGFQHG-- 146
DB 422 TALGNLSMKA--KAAADTDGALGEGI--SKDOEVKAGETVTFKAGKMLKVKODGANFTYSIQ 478
OY 147 DAFTRMLKIP-----ENKQLFTEKNIIIL-----GIDWTDENVIPTRGYINLTYK 192
DB 479 DALTGLTSTLTGCTTNGGNDATVINKDGLTTPAGNGCTGTNTISVTKDGIRKAGNKAI 538
OY 193 TEEAGWLAGYANASFLAKKFPSPDKRRAIYVGGISPAVTFIAGIYLAGIKAMNLSKD 252
DB 539 TNVAGSLRAYDANF-----DVLNNSATDINRHVEDA-----YKGLNLNLEKKNAN 583
OY 253 KTKITTTDKIEI-----NLGFDVQ--DTSTKERLEQIASKDKPSTLLAVAGPLTEIFSII 306
DB 544 KQPLVTDSTAAITVGLRLKIGWVSTKNGTKESNOYKQADE--VLEFTGAGATVTSK--- 638
OY 307 ANQNDRIYL-----GVDTDQSLVYVTKTKNKFSTSLK--NLGYSVPSVLSDLYT 353
DB 639 -SENCKHHTIVASVETKADCGLEKDGDTIKLKVNDNQNTDNLVTGNNNGAV-----T 689
OY 354 KKSNSRNLAGFEFGKSKATVYIGIDR--FVDIADTSLGENDKKLITAEISEKKKEFEET 412
DB 690 K-----GGEFTVKTGAT-----DADRGAVPVKADATANDADKKVAT--VKDVATAINSAA 736
OY 413 KTIPEAVRKTLEIPMPDKOPDKQOESLKLITDINKN 451
DB 737 TFEVKTENLTSTIDENPTDNGKDALKAGDITLTFKAGKN 775

```

RESULT 14

US-08-685-467-4

; Sequence 4, Application US/08685467

; Patent No. 6060059

; GENERAL INFORMATION:

; APPLICANT: St. Gene III, Joseph W.

; APPLICANT: Barenkamp, Stephen J.

; TITLE OF INVENTION: HAEMOPHILUS ADHESION PROTEINS

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert

; STREET: Four Embarcadero Center, Suite 3400

; CITY: San Francisco

; STATE: California

; COUNTRY: United States

; ZIP: 94111-4187

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/685,467

; FILING DATE: 22-JUL-1996

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/409,995

; FILING DATE: 24-MAR-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Silva, Robin M.

; REGISTRATION NUMBER: 38,304

; REFERENCE/DOCKET NUMBER: A-61053-2/RFT/RMS

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 781-1989

; TELEFAX: (415) 398-3249

; TELEX: 910 277299

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1912 amino acids

; TYPE: amino acid

; STRANDEDNESS: unknown

; TOPOLOGY: unknown

; MOLECULE TYPE: protein

; US-08-685-467-4

Query Match 5.1%; Score 117; DB 3; Length 1912;

Best Local Similarity 21.0%; Pred. No. 0.2;

Matches 109; Conservative 65; Mismatches 217; Indels 128; Gaps 24;

```

OY 12 GIGLVFPLSATATTISAGCWDKETEKEKSADNQNQI-----TDVSKISG-----LVN 59
DB 306 GNGLTATAKAVIDAIVKAKGRVKT-----TANGQNGDFATVAGSTVTEPSGGTTASVTK 361
OY 60 EKKSSEIMAKADANKHFGILN-----AIVTAGTVND----- 91
DB 362 DTNGGITYVKKADAKGDDGKFPDSDKKIYADTTALTVTGGKVAEIAKEDDKKLVNAGDLY 421
OY 92 NSFNOSSWEAIDQALNGEITTSVDSSTAELGKYSLSLANTNKNVWYL---SGFQHG-- 146
DB 422 TALGNLSMKA--KAAADTDGALGEGI--SKDOEVKAGETVTFKAGKMLKVKODGANFTYSIQ 478
OY 147 DAFTRMLKIP-----ENKQLFTEKNIIIL-----GIDWTDENVIPTRGYINLTYK 192
DB 479 DALTGLTSTLTGCTTNGGNDATVINKDGLTTPAGNGCTGTNTISVTKDGIRKAGNKAI 538
OY 193 TEEAGWLAGYANASFLAKKFPSPDKRRAIYVGGISPAVTFIAGIYLAGIKAMNLSKD 252
DB 539 TNVAGSLRAYDANF-----DVLNNSATDINRHVEDA-----YKGLNLNLEKKNAN 583
OY 253 KTKITTTDKIEI-----NLGFDVQ--DTSTKERLEQIASKDKPSTLLAVAGPLTEIFSII 306

```



```
Db 584 KQPLVYDSTAAVGDRLKLGWVSTKNGTKEESNOVKQADE--VLEFGAGAAVTYSK--- 638
OY 307 ANONDRYLI-----GVDTDOSLYVTKTKNKEFTSILK--NLGYSVFVLSLDLYT 353
Db 639 -SENGKHTTIVSAEYKADCGLEKDDGTILKLYKNDONTNVLVGNNGTAV-----T 689
OY 354 KKSNSRNLAGEFEGKKSATVYLGIKDR-FVADIADTSLEGNDRKLATPAISAAKKEFEKT 412
Db 690 K-----GGFETVKTGAT----DADRGKVTYKDATANDADKKVAT--VKDVATAINSAA 736
OY 413 KTIPEAEVKTLEIPEKPKOPKQOESLDKLTIDINKN 451
Db 737 TFVKTEMLTTSIDEDNPTDNGKDALKAGDTLTFKAGKN 775
```

RESULT 15

```
US-09-377-155-33
; Sequence 33, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
; APPLICANT: PEAK, Ian Richard Anselm
; APPLICANT: JENNINGS, Michael Paul
; APPLICANT: MOXON, E. Richard
; TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
; FILE REFERENCE: 065064/0128
; CURRENT APPLICATION NUMBER: US/09/377,155
; CURRENT FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: PCT/AU98/01031
; PRIOR FILING DATE: 1998-12-14
; PRIOR APPLICATION NUMBER: GB 9726398.2
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 33
; LENGTH: 2353
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
; US-09-377-155-33
```

Query Match 5.1%; Score 117; DB 4; Length 2353;

Best Local Similarity 21.0%; Pred. No. 0.28;

Matches 109; Conservative 65; Mismatches 217; Indels 128; Gaps 24;

```
OY 12 GLGLVPLSAITISAGCWMKETTKEESADNONKQI-----TDVSKISG-----LVN 59
Db 307 GNGLVYAKAVIDAVNKAQWVYKTT---TANGQGDPAIVASGTNVFESGDGTTASVTK 362
OY 60 ERKSEIIMAAKADANKHEGLNM-----AIVTAGTAVND----- 91
Db 363 DTNGNGITVYDAKVGGLGKLFDSDKIVADTTLTYTGKVAELAKEDDKKLVNAGDLV 422
OY 92 NSFNQSWEAITQOLGALTGCEITSVDSSTAELEGKYSILANTKNVYL--SGFQHG-- 146
Db 423 TALGNLSWKA--KAADDTGDALEGI--SKDEYVAGETVTFKAGKNLKVODGANFTYSLO 479
OY 147 DAFTRMLKIP-----ENKOLFTEKNIIIL-----GIDWTDENVIPTRINILYK 192
Db 480 DALGLTSLTYLGTNGNGNDKATVINKDGLTTPACNGCGTTGTNTISVTKGIRAKGNKAI 539
OY 193 TEEAGWLAGYANASFLAKKFPSPDKRSALVIGGISPATDFLAGYLAGIKAMNLKNSD 252
Db 540 TNVASGLRAYDDANF-----DVLNNSATDLNRHVEDA-----YKGLNLNNEKNAN 584
OY 253 KTKKITTDKLEI-----NLGFDVQ--DTSTKERLEQIASKDKPSTLLAVAGPLTEIFSDII 306
Db 585 KQPLVYDSTAAVGDRLKLGWVSTKNGTKEESNOVKQADE--VLEFGAGAAVTYSK--- 639
OY 307 ANONDRYLI-----GVDTDOSLYVTKTKNKEFTSILK--NLGYSVFVLSLDLYT 353
Db 640 -SENGKHTTIVSAEYKADCGLEKDDGTILKLYKNDONTNVLVGNNGTAV-----T 690
```

```
OY 354 KKSNSRNLAGEFEGKKSATVYLGIKDR-FVADIADTSLEGNDRKLATPAISAAKKEFEKT 412
Db 691 K-----GGFETVKTGAT----DADRGKVTYKDATANDADKKVAT--VKDVATAINSAA 737
OY 413 KTIPEAEVKTLEIPEKPKOPKQOESLDKLTIDINKN 451
Db 738 TFVKTEMLTTSIDEDNPTDNGKDALKAGDTLTFKAGKN 776
```

Search completed: July 18, 2002, 09:14:54
Job time: 34 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2002, 15:29:13 ; Search time 50.25 seconds
(without alignments)
1552.651 Million cell updates/sec

Title: US-09-676-249A-2
Perfect score: 2299
Sequence: 1 MKRRIKMKFLGLVPEPLS.....KQPDKQESLDKLTIDINKN 451

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Minimum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP RODENT:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*
15: SP_VIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEP:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	706	30.7	457	2	048902 mycoplasma
2	554.5	24.1	461	16	0980L5 mycoplasma
3	417	18.1	465	2	09X775 mycoplasma
4	400.5	17.4	428	2	052311 mycoplasma
5	397.5	17.3	428	2	09RGX5 mycoplasma
6	394.5	17.2	428	2	032417 mycoplasma
7	394	17.1	429	2	09RGX6 mycoplasma
8	394	17.1	429	2	09RGX4 mycoplasma
9	393.5	17.1	428	2	09RGX7 mycoplasma
10	389.5	16.9	428	2	09RGX3 mycoplasma
11	387.5	16.9	428	2	09RGX6 mycoplasma
12	332.5	10.1	350	16	09RGX4 mycoplasma
13	229.5	10.0	349	2	09EY91 streptococc
14	223	9.7	350	16	097RH0 streptococc
15	207	9.0	350	16	09CFM9 lactococcus
16	195.5	8.5	357	16	097L60 clostridium

17	187.5	8.2	516	16	09PRD3 ureaplasma
18	183	8.0	342	16	09A1P7 streptococc
19	183	8.0	348	2	09AK41 streptococc
20	181.5	7.9	359	16	09XKV7 thermotoga
21	174	7.6	357	16	092BW7 listeria in
22	172	7.5	341	2	09EPD3 borrelia in
23	171	7.4	264	2	031358 borrelia ga
24	168.5	7.3	525	16	09POR5 ureaplasma
25	165	7.2	347	2	09AK42 streptococc
26	164	7.1	524	16	09PRC9 ureaplasma
27	162	7.0	339	2	09S6C1 borrelia bu
28	162	7.0	339	2	087960 borrelia bu
29	161	7.0	335	2	09R776 borrelia bu
30	158	6.9	339	2	09EPD2 borrelia bu
31	157	6.8	325	2	09R777 borrelia bu
32	149	6.5	329	16	09RGW4 rhizobium l
33	146	6.4	264	2	031359 borrelia ga
34	144.5	6.3	243	2	031361 borrelia ga
35	144	6.3	334	2	093V09 borrelia bu
36	143.5	6.2	383	2	046239 clostridium
37	141.5	6.2	539	16	09P010 ureaplasma
38	141.5	6.2	889	2	09L961 moraxella c
39	141	6.1	379	16	09RSQ4 deinococcus
40	136.5	5.9	912	2	09XD52 moraxella c
41	133.5	5.8	330	16	0926D6 rhizobium m
42	131.5	5.7	1786	5	09U0P0 plasmodium
43	130	5.7	484	16	098RB2 mycoplasma
44	130	5.7	894	2	09L962 moraxella c
45	130	5.7	1170	2	09AJR8 erysipeloct

ALIGNMENTS

RESULT	ID	PRELIMINARY	PRT	457 AA.
048902	048902	048902		
AC	01-NOV-1996 (TREMBLrel. 01, Created)			
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
DE	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
OS	Ag 243-5 PROTEIN PRECURSOR.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Molluscites;			
OX	NCBI_TaxID=2094;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE-96163149; PubMed=8551970;			
RA	Ushio S., Iwaki K., Tanai M., Ohta T., Fukuda S., Sugimura K.,			
RA	Kurimoto M.;			
RT	"Metastatic promoting activity of a novel molecule, Ag 243-5 derived			
RT	from Mycoplasma and the determination of the complete nucleotide			
RT	sequence."			
RL	Microbiol. Immunol. 39:393-400(1995).			
DR	EMBL; D16674; BAA04082.1; -			
DR	InterPro: IPR003760; Bmp.			
DR	Pfam: PF02608; Bmp; 1.			
KW	Signal.			
FT	SIGNAL 1 25			
FT	CHAIN 26 457			
FT	POTENTIAL. AG 243-5 PROTEIN.			
SO	SEQUENCE 457 AA; 50789 MW; 9AEEB11620CB22F CRC64;			

Query Match 30.7%; Score 706; DB 2; Length 457;
Best Local Similarity 38.4%; Pred. No. 8.7e-33;
Matches 173; Conservative 87; Mismatches 145; Indels 46; Gaps 16;

OY	14	GLVPLSAIAIT-ISAGCWDKETTKEESADNQNIIDVSKISGLVNERSEIMAAKADA 72
DB	12	GTISTVASVATFVSCG---ETDKGKII-----RFD-----NSFVDRQAEIKAKA---- 54
OY	73	NKHEGLMAIVTAGTVDNSFNQSSWEAI---QALGALTG-GEITSVDSFALEBGKY 127

```

Db 55 -NPFENVLLTAGTGVOQKSFNQSIWEAVLEHYDQIEKTYMLDRVSOETNNQSELIGKY 113
Qy 128 SSIANTNNKNNWVLSGFQHDGDAFTRWLIKIPEN-----KQFTEKNIIIGIDW---TDTEN 179
Db 114 KRFPLNGKNNWVLLTGQOQOEFPKFKQTDNSGKKSSDLAEKRVITVAVNDMLSKEDKD 173
Qy 180 VIPTGRYINLYKTEEAGNLGAVANASFLAKKPPSDPTRRSALVIGGISPAVTEFIAGY 239
Db 174 LKAGHFISLTKTEEBAGFIAGYASSKFLAYKFPNDKARTIAPFCGAGACAGATDFIAGF 233
Qy 240 LAGIGAMNLANSDKTKITTDKIEINLGFVDVDTSTKELEQIASDKRSTLLAVAGPT 299
Db 234 LAGIAKNNNDNPJAKVYITSDNNINIDTGF-ISNDKTATEFINGIVKNS--SLVLPVAGSLT 290
Qy 300 EIFSDII--ANONDRYLIGVPTDOSLVYTKTKNKPFTSTILKNLGYSVFSLDLYTKKSN 357
Db 291 SSVVDALKSNNDTKLIVDTDOSKIFSPA-TVFTSTIEKHLGRITIVYLDIMLKEDD 349
Qy 358 SBNLAGFEFGK---KSATVYLGIKDRFYDIADTSLGNDKRLATEAISAKKEFEKTK 413
Db 350 SKFLGSFRSFKLTPNANATVYKISDDEFGVGSNSTVADADKYKAGDFLEATADAFKKQIQ 409
Qy 414 TIPAEVETKLEIPEK---PDKQPDKQDESL 441
Db 410 ANPT-NYKSVLGIPTMLINDNDKDKNEKASL 439

```

```

RESULT 2
ID 0980L5 PRELIMINARY: PRT: 461 AA.
AC 0980L5:
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE ABC TRANSPORTER XYLOSE-BINDING LIPOPROTEIN.
GN MYPU_3460.
OS Mycoplasma pulmonis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2107;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UAB CT1P;
RX MEDLINE=21267165; PubMed=11353084;
Chambaud I., Dyrbylg K., Wroblewski H., Viaril A., Rocha E.P.C.,
Moszer I., Dybvig K.,
Blanchard A.;
"The complete genome sequence of the murine respiratory pathogen
RT Mycoplasma pulmonis."
RL Nucleic Acids Res. 29:2145-2153(2001).
DR EMBL: AL445564; CAC13519.1; -.
DR Mypulist: MYPU_3460; -.
DR InterPro: IPR003760; Bmp.
DR Pfam: PF02608; Bmp; 1.
KW Complete proteome.
SQ SEQUENCE 461 AA; 51096 MM; 6724D4D820809CE4 CRC64;

```

```

Query Match 24.1%; Score 554.5; DB 16; Length 461;
Best Local Similarity 30.3%; Pred. No. 4.1e-24;
Matches 140; Conservative 88; Mismatches 175; Indels 59; Gaps 13;
Qy 5 IKNNFLGIGLVFPPLSATATISAGCWDKETTKEKSDNONKOIT--DVSRTISGLVNER- 61
Db 1 KMLNK-KLFSTLFPVAAIALPATF-----VSCAONPKNTNNSDSSKTTDLISQKE 51
Qy 62 -----KSEITMAKADANKHFGMLMAIYTAGTVNDNSFNQSSWEAIIQGLGALTGERI 113
Db 52 VTETQKIVENKIKQASLETFQK-----VYLIPADGNIDDKSFNQGYEESQKTLAKDPVDKAY 106
Qy 114 TS-----VDSSTALEGKYSSSLANTNNKNNWVLSGFQHDGDAFTRWLIKIPENK 159

```

```

Db 107 KSQNEEAENOHLKNDYNSAVKDLQNKVVALDRGYTTWILTGFOQGEIENFLDENDNL 166
Qy 160 QLFTEKNIIIGIDWTDTEEN-VIPTGRYINLYKTEEAGNLGAVANASFLAKKPPSDPTK 218
Db 167 RRFKEKRVKVIIGVDNAPNANSKIPGOSLSLFLKTEEBAGMAGYASADFLGRTKYANNEAK 226
Qy 219 RSAIVIGGISPAVTEFIAGYLAGIKNAN--LKNSDKTKITTDKIEINLGFVDVDTSTK 276
Db 227 RAISAFGGDGFAGVNDLFGFEGGIRAWNSBAENKKNKVIYSEMLVLDTF-IPNAEKN 285
Qy 277 ERLEQIASDKRSTLLAVAGPTLEIFSDII--ANONDRYLIGVPTDOSLVYTKTKNKF 333
Db 286 EVVSNVETGKSTISLPAVAGPTGVVVDLRRDTSDEDERFIYVDTDOSLSTFNDSKRFF 345
Qy 334 TSLIKNLGYSVFSLDLYT-----KSNRNLAGFEFGKSGATVYLGIKDRFYDIA 385
Db 346 TSVINIAFPVQIILLATLKDESVYIAKEGDKFLGS---NPKMLVLKRLGISAKFVNIT 402
Qy 386 DTSLGNDKRLATEAISEA--KKEFEKTKTIPAEVETKLE 425
Db 403 KSRVKEISIKTQADTSIQKAIKDKMANPNNSKRIEKMVTNGDLE 444

```

```

RESULT 3
ID 09X775 PRELIMINARY: PRT: 465 AA.
AC 09X775:
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE P48 MEMBRANE LIPOPROTEIN PRECURSOR.
GN P48.
OS Mycoplasma agalactiae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2110;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M7;
RX MEDLINE=20002620; PubMed=10531294;
RA Rosatti S., Pozzi S., Robino P., Montinaro B., Conti A., Fadda M.,
Pittau M.;
"P48 major surface antigen of Mycoplasma agalactiae is homolog to a
RT malp product of Mycoplasma fermentans and belongs to a selected family
RT of bacterial lipoproteins."
RL Infect. Immun. 67:6213-6216(1999).
DR EMBL: AJ132423; CAB43718.1; -.
DR InterPro: IPR003760; Bmp.
DR Pfam: PF02608; Bmp; 1.
KW Signal; Lipoprotein.
FT SIGNAL 1
FT CHAIN 22
FT 23 465 P48 MEMBRANE LIPOPROTEIN.
SQ SEQUENCE 465 AA; 51149 MM; 60AD5448CFE03C6 CRC64;

```

```

Query Match 18.1%; Score 417; DB 2; Length 465;
Best Local Similarity 28.4%; Pred. No. 3.1e-16;
Matches 141; Conservative 94; Mismatches 180; Indels 82; Gaps 22;
Qy 3 KIKKNNKFLGIGLVFPPLSATATISAGCWDKETTKEKSDNONKOITDVSRTISGL-VNER 61
Db 2 KKNKFLFLGAA--PVLSPVLVAASCGDKYFKETEVGVGVTISTILAHITSRKGLKEEG 58
Qy 62 KSEITMAKADANKHFGMLMAIYTAGTVNDNSFNQSSWEAIIQGLGALTGSDSTA 121
Db 59 LTVENAPKA-----TFTTDEGSVHDSFNSGWEAVHAKVYELGLDKRAOV-SGNK 107
Qy 122 ELEGR-----YSSLANTNNKNNWVLSGFQHDGDAFTRWLIKIPENK-OLFTEKNII 168
Db 108 NLRNKVPKPKGQLLEAKNMIIDSGFRIVYIGCFHQASL---VGLDENVIKKIKIDNNII 164
Qy 169 ILGIDWT-----DTENV-----IPTGRYINLYKTEEAGNLGAVANASFLAKKPPSDPTK 218

```

```

Db      165 FTTVDENFETEDDANKVEIKKIGEGHLYPVLTFRQKAAYIAGRALADYFSQVYKDQPK 224
Oy      219 KSAIVGGISPAVTDFEITAGYLGIKAMNMLKNSDKRKITYTDKEINLGEVDOTSTKER 278
       | : ||| ||||| ||| ||| : ||| : ||| : ||| :
Db      225 RTGAFGAPGWPAWDEFTAGTFOGLIDMWKHEPPEAKTKSLNETIELNTLF---TSQTQA 281
Oy      279 LEOIAKNOKPSLLAVAGPL-TEITSDI--IANONDRYLIGVDGOSNLVYTKN----- 330
       | | | | | : ||| | : : | : | : ||| ||| |||
Db      282 TTAISVVKATASYPVAGSLSTDTPAKETKKLADK-DKFLTGADAO-----KNALKGH 333
Oy      331 KEETSILKNLGYSVESVLSIDLTYTKSNSRNL-AGEFEGRKSAT-VYLGIKD---REYDI 384
       | : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db      334 RIETSVMKRIQGAAYNIILADLYSKSGNQDLDPGEIGKKNGTPTVFYGDTEDMQYGV 393
Oy      385 ADTS-LEGDNKLAEALSEAKKEPEEKRTTPAEVKTLEI-----PEMPDKQ 433
       | : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db      394 ATSGILDLDNDINDAIKADATAYVQK-KT---EIQSLKDOMEYAKKALGARFPSPD 448
Oy      434 PDKOOESLDKLTITDINK 450
       | : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db      449 GGQFGKMVDWLASETRK 465
SUBMIT 4
ID      052311 PRELIMINARY; PRY; 428 AA.
AC      052311:
DT      01-JUN-1998 (TREMBLrel. 06, Created)
DT      01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT      01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE      MEMBRANE LIPOPROTEIN P48V.
OS      Mycoplasma fermentans.
OC      Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC      Mycoplasmataceae; Mycoplasma.
OX      NCBI_TextID=2115;
RN      [1]
RC      SEQUENCE FROM N.A.
RA      STRAIN-PG18;
RW      Rawadi G., Dyer K., Dujeancourt A.;
RL      Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RR      EMBL: AF036106; AAB89740.1; -.
DR      InterPro: IPR003760; Bmp.
DR      Pfam: PF02608; Bmp; 1.
KW      Lipoprotein.
SQ      SEQUENCE 428 AA; 47862 MW; D9506BE817E330EEA CRC64;
```

```

Query Match      17.4%: Score 400.5; DB 2; Length 428;
Best Local Similarity 27.3%: Pred. No. 2.4e-15;
Matches 129; Conservative 95; Mismatches 175; Indels 73; Gaps 21,
QY      5 IKWNFFLGIGLVPEISAI-ATISAGCWKETTKREKSADNOKOI--IDVKSISGLVNER 61
      1 MKRSKKILIGT-SPFAAILPAVAVC-----KNDNESISPEREKDISKRYTTTNGANG 50
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      62 KSEIMAAKADAKHHGILNMAIYTAGTVDNDSFNQSSWEAIOQLATJGELTTSVDSSTA 121
      51 KOVVAEAL-----LKLRPVLTIDBGSKIDDKSFNSAFEPALAIKQYIEINSEVPS-S 104
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      122 ELEGKYSSILANTNKNWVVLGSGFHODATFRMLKIPENKOLFTEKNNILIGIDMTQENVI 181
      105 NFESAIYNSALSGHKIHWLNGFHKHOOSIKOY--IDAHRELEDRNOIKITIGDF-DIELEY 161
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      183 PTGRYINILTYKTEEGWAGLYANASAFIAKKFPSPDPTKRAAIYIGGISPAYVDFTAGYLA 241
      162 KM--FYSLQFNKRESAFTTGYALASWLSQ---DESKRVASAFSGGAFPGVTTTFDEGFAK 216
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      242 GIKANNLKNDSKRTKIT-TDKIEMNGEVDVDTSTKERLEQIAS-----KDKPST 290
      217 GLIYYLNQKH--KSVAFETHHSVPYKLDGSG-----TAGEKNNYIINNVYSTPADVYKNPBY 269
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      291 LIAVAGPLETIRSDIIANONRILYIGVDTSGLVYTKTKNKKFTSILKNLGVSVFVSVD 350
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Db      270 ILISVAGPAT--FETVYKLANKGQVYIGVDSQGM--QDKRILISVYLNHIQAVETILLD 325
Qy      351 L-----YTKKSNSNRNLNLAFFEGGKSAATVYLIGRIDREVDIADTSLGNDKKL-ATE 399
Db      326 LILLEKEGKPYVVDKKADKKKSHFGQ-----KEKIVGAENFNPSTEDAKINN 377
Qy      400 AISEAKKEFEETKTIPAEVYKTLLEIPMD-KQPKQOESLDKLTIDINK 450
Db      378 KIKEAKIKMKPE---LPEDFVYKINSRKALKGKNKIDNVSRLEALISAIKK 425

RESULT      5
Q9RGX5      Q9RGX5      PRELIMINARY;      PRT;      428 AA.
AC      Q9RGX5;
DT      01-MAY-2000 (TREMBLrel. 13, Created)
DT      01-MAY-2000 (TREMBLrel. 13, last sequence update)
DT      01-JUN-2001 (TREMBLrel. 17, last annotation update)
DE      MACROPHAGE ACTIVATING LIPOPROTEIN-404 PRECURSOR.
GN      MALP.
OS      Mycoplasma fermentans.
OC      Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC      Mycoplasmataceae; Mycoplasma.
OX      NCBI_TaxID=2115;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-SK5;
RC      MEDLINE=99115554; PubMed=9916088;
RA      Calcutt M.J., Kim M.F., Karpas A.B., Muhlradt P.F., Wise K.S.;
RT      "Differential posttranslational processing confers intraspecies
RT      variation of a major surface lipoprotein and a macrophage-activating
RT      lipopeptide of Mycoplasma fermentans.";
RL      Infect. Immun. 67:760-771(1999).
DR      EMBL; AF099211; AAD16395.1; -.
DR      InterPro; IPR003760; Bmp.
DR      Pfam; PF02608; Bmp; 1.
KW      Signal; Lipoprotein.
FT      SIGNAL      1      24      POTENTIAL.
FT      CHAIN      25      428      MACROPHAGE ACTIVATING LIPOPROTEIN-404.
SQ      SEQUENCE      428 AA;      47835 MM;      DO3F047RA2B1460 CAC64;

```

[illegible]

QY 402 SEAKKEFEKTKTTPAEVYRKTLEIPEMPD-KOPDKQOESLTKITDINK 450
 DB 380 KEAIMEKFE---LPEDFVKYINSOKALKDKGNKIDNVSERLAIIAISAIK 425

RESULT 6
 ID 032417 PRELIMINARY; PRT; 428 AA.
 AC 032417;
 DT 01-JAN-1998 (TREMblrel. 05, Created)
 DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE M16AG.
 OS Mycoplasma fermentans.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
 CC Mycoplasmataceae; Mycoplasma.
 NX NCBI_TaxID=2115;
 [1]
 SEQUENCE FROM N.A.
 TISSUE=NONE MARROW;
 RX MEDLINE=98022661; PubMed=9359703;
 RA Matsunoto M., Takeda J., Inoue N., Hara T., Hatanaka M., Takahashi K.,
 RA Nagasawa S., Akeo H., Seya T.;
 RT "A novel protein that participates in nonself discrimination of
 RT malignant cells by homologous complement.";
 RL Nat. Med. 3:1266-1270(1997).
 RN [2]
 SEQUENCE FROM N.A.
 RA Matsunoto M., Seya T.;
 RL Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 SEQUENCE FROM N.A.
 RX MEDLINE=98241611; PubMed=9575196;
 RA Matsunoto M., Nishiguchi M., Kikkawa S., Nishimura H., Nagasawa S.,
 RA Seya T.;
 RT "Structural and functional properties of complement-activating protein
 RT M16Ag, a Mycoplasma fermentans gene product that induces cytokine
 RT production by human monocytes.";
 RL J. Biol. Chem. 273:12407-12414(1998).
 RN [4]
 SEQUENCE FROM N.A.
 RA Seya T., Matsunoto M.;
 RT "MEMBRANOUS PROTEIN M16AG AND CYCLIC-DNA CAPABLE OF CODING THE
 RT SAME.";
 RL Patent number JP1997157295-A/1,
 RL 17-JUN-1997. KAKAKU GIJUTSU SHINKO JICYODAN.
 DR EMBL; D64083; BAA23530.1; -
 DR InterPro; IPR003760; Bmp.
 DR Pfam; PF02608; Bmp; 1.
 SQ SEQUENCE 428 AA; 47862 MW; F43B078F21DADD6E CRC64;

Query Match 17.2%; Score 394.5; DB 2; Length 428;
 Best Local Similarity 26.8%; Pred. No. 5.4e-15;
 Matches 126; Conservative 101; Mismatches 174; Indels 69; Gaps 20;

QY 5 IKMNFELGLVFPPLSAI-ATISAGCMDEKTTKEKSDONKOTIDVSKISGLVNERKS 63
 DB 1 MKSKKILGL-SPIAAILPVAVSCGNDES-----NISFEKDKISKYTTTANGKO 52

QY 64 EIMAKADANKHFGIMAVTAGTVNDNSFNQSSWEAIOOLGALTGEITSVDSSTAEI 123
 DB 53 VVKNEL-----LKLKPLVITDEKIDKSFNQSFAELKAIKNOTGIEINNVERP-SNF 106

QY 124 EGRYSILANTKNVWVLSGFQHGDAFTRMKLPENKOLFTEKNIIILIGIDWTDTEENVPT 183
 DB 107 ESAYNSALSAGHKIIVLNGFKHQOSIKOY--IDAHRRELENNQIKITIGIDF-DIETEKW 163

QY 184 GRVILYTKTEAGWLAGYANASFLAKKPPSDPTRSAIVIGGGISPAVTFPIAGYLAGI 243
 DB 164 --FYSLOFNKESAFETTYGAIASWLSQ---DESKRVVASEFGGAFPGVTFTEGFAKI 218

QY 244 KAMLNKSDKTKI-TTDKIEINLGFVDYSTKERLEQIAS-----KDKPSTLL 292
 DB 219 LYNNQKH--KSKYIHTSPVKIDSGF-----TAGKMTVINNVLSSTPADVKYVPHYL 271

QY 293 AVAGPLEIFSDIINQNDRIILGVDTSQSLVYTKTKKFTSIIKNGIVSVFSLD- 351
 DB 272 SVAGPAT--FETVRLANKQYVIGVSDQGM--ODKRIITSVLKHKAQVETLADI 327

QY 352 -----YTKSNSRNLAGFEGKKSATVYLAGIKDRFVDIADTSLGNPKKL-ATPAI 401
 DB 328 LEKEGYKPYVVKDKKAKKMSHFGTQ-----KEWIGVAENHFNTEGEQAKINKKI 379

QY 402 SEAKKEFEKTKTTPAEVYRKTLEIPEMPD-KOPDKQOESLTKITDINK 450
 DB 380 KEAIMEKFE---LPEDFVKYINSOKALKDKGNKIDNVSERLAIIAISAIK 425

RESULT 7
 ID 09RGX6 PRELIMINARY; PRT; 429 AA.
 AC 09RGX6;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE MACROPHAGE ACTIVATING LIPOPROTEIN-404 PRECURSOR.
 GN MALP.
 OS Mycoplasma incognitus.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
 CC Mycoplasmataceae; Mycoplasma.
 NX NCBI_TaxID=2101;
 [1]
 SEQUENCE FROM N.A.
 RC STRAIN=INCIGNITUS;
 RX MEDLINE=99115554; PubMed=9916088;
 RA Calcutt M.J., Kim M.F., Karpas A.B., Muhradt P.F., Wise K.S.;
 RT "Differential posttranslational processing confers intraspecies
 RT variation of a major surface lipoprotein and a macrophage-activating
 RT lipopeptide of Mycoplasma fermentans.";
 RL Infect. Immun. 67:760-771(1999).
 DR EMBL; AF099210; AAD16394.1; -
 DR InterPro; IPR003760; Bmp.
 DR Pfam; PF02608; Bmp; 1.
 KW Signal; Lipoprotein.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 429 MACROPHAGE ACTIVATING LIPOPROTEIN-404.
 SQ SEQUENCE 429 AA; 47961 MW; B6D09A8B12AC3171 CRC64;

Query Match 17.1%; Score 394; DB 2; Length 429;
 Best Local Similarity 27.0%; Pred. No. 5.8e-15;
 Matches 128; Conservative 101; Mismatches 169; Indels 76; Gaps 21;

QY 5 IKMNFELGLVFPPLSAI-ATISAGCMDEKTTKEKSDONKOTIDVSKISGLVNERKS 63
 DB 1 MKSKKILGL-SPIAAILPVAVSCGNDES-----NISFEKDKISKYTTTANGKO 52

QY 64 EIMAKADANKHFGIMAVTAGTVNDNSFNQSSWEAIOOLGALTGEITSVDSSTAEI 123
 DB 53 VVKNEL-----LKLKPLVITDEKIDKSFNQSFAELKAIKNOTGIEINNVERP-SNF 106

QY 124 EGRYSILANTKNVWVLSGFQHGDAFTRMKLPENKOLFTEKNIIILIGIDWTDTEENVPT 183
 DB 107 ESAYNSALSAGHKIIVLNGFKHQOSIKOY--IDAHRRELENNQIKITIGIDF-DIETEKW 163

QY 184 GRVILYTKTEAGWLAGYANASFLAKKPPSDPTRSAIVIGGGISPAVTFPIAGYLAGI 243
 DB 164 --FYSLOFNKESAFETTYGAIASWLSQ---DESKRVVASEFGGAFPGVTFTEGFAKI 218

QY 244 KAMLNKSDKTKI-TTDKIEINLGFVDYSTKERLEQIAS-----KDKPSTLL 292
 DB 219 LYNNQKH--KSKYIHTSPVKIDSGF-----TAGKMTVINNVLSSTPADVKYVPHYL 271

```

OY 293 AVAGPTEIFSDI-IANODRVLIGVDTOOLVYTKNKFPTSLKNGSVSVLSL 351
DB 272 SVAGPAT--FETVRLAANGQYIVGDSOGMT--ODKDRILTSVLKHIOQAVYETLLD 327
OY 352 -----YTKKSNSRLAGFEFGKKSATVYLGIDKDFVDIADTSLGSDNKL-ATEA 400
DB 328 ILEKEGYPYVVKDKKADKMSHFQTQ-----KEWIGVAENHNSFTEQOAKINNK 379
OY 401 ISEAKKEF-----EKTKTIPAEVRKTLPEMPDKOPKOQESLDKLTIDINK 450
DB 380 IKAIMKFELEDPEFVKYINSOKVLK-----DGKNKIDVSRERLALITSAINK 426

RESULT 8
O9RGX4 PRELIMINARY: PRT: 429 AA.
AC 09RGX4:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE MACROPHAGE ACTIVATING LIPOPROTEIN-404 PRECURSOR.
MALP.
Mycoplasma fermentans.
Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
Mycoplasmataceae; Mycoplasma.
NCBI_TaxID=2115;
[1]
SEQUENCE FROM N.A.
RP STRAIN=MT-2;
RC MEDLINE=99115554; PubMed=9916088;
RX Calcutt M.J., Kim M.F., Karpas A.B., Muhlradt P.F., Wise K.S.;
RA "Differential posttranslational processing confers intraspecies
RT variation of a major surface lipoprotein and a macrophage-activating
RT lipopeptide of Mycoplasma fermentans.";
RL Infect. Immun. 67:760-771(1999).
DR EMBL; AF099212; AAD16393.1; -.
DR InterPro; IPR003760; Bmp.
DR Pfam; PF02608; Bmp; 1.
KW Signal; Lipoprotein.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 429 MACROPHAGE ACTIVATING LIPOPROTEIN-404.
SQ SEQUENCE 429 AA; 47933 MW; B6D08CF975AC3171 CRC64;

```

```

Query Match 17.1%; Score 394; DB 2; Length 429;
Best Local Similarity 27.0%; Pred. No. 5.8e-15;
Matches 127; Conservative 102; Mismatches 172; Indels 70; Gaps 21;

```

```

OY 5 IWKNKFLGLVFPISAI-ATISAGCWDKETTKEKSADNQNKOITPVSKISGLVNERKS 63
DB 1 MKRSKRLILGL-SPIALILPAVAVSCGNDES-----NISFEKDISKTYTTNANGKQ 52
OY 64 EIMAAADANKHGLMAIYTAGVYNDNSFNQSSWEALIOQLALNGEITTSVDSSTAE 123
DB 53 VVKNAL-----LKLKPVLTDEGKIDKSFNQSAPFALKAIKQVIEINNEPS-SNP 106
OY 124 EGRYSILANTNKNWVYLSGFQHGDAFTRMUKIPEKNOLFTEKNIILIGIMTWTENV 183
DB 107 ESYNALSAGHKIWLNGFKHQOSIKQY--IDAHRERLERNQIKIKGIDF-DIEIEYKW 163
OY 184 GRVINTLYTEAGWLAGYANASFLAKFPSPDTPKRSATVYIGGISPAVDTFAGYLAG 243
DB 164 --FYSLOFINKESEFTTGVAIASWLSQ---DESKRVASFGGAPFGVTTFEGRAKGI 218
OY 244 KAMNLSKDKRTKI-TTDKIEINLGFVDOTSTKERLEQIAS-----KDKSSTLL 292
DB 219 LYYNQNH--KSKKIYHTSPVKLDSG-----TAGCKMNVYINNVLSSTPADVYVNPVIL 271
OY 293 AVAGPTEIFSDI-IANODRVLIGVDTOOLVYTKNKFPTSLKNGSVSVLSL 351
DB 272 SVAGPAT--FETVRLAANGQYIVGDSOGMT--ODKDRILTSVLKHIOQAVYETLLD 327
OY 352 -----YTKKSNSRLAGFEFGKKSATVYLGIDKDFVDIADTSLGSDNKL-ATEA 400

```

```

DB 328 ILEKEGYPYVVKDKKADKMSHFQTQ-----KEWIGVAENHNSFTEQOAKINNK 379
OY 401 ISEAKKEF-----EKTKTIPAEVRKTLPEMPDKOPKOQESLDKLTIDINK 450
DB 380 IKAIMKFE-----LPEDFVKYINSOKALDKGNKIDVSRERLALITSAINK 426

RESULT 9
O9RGX7 PRELIMINARY: PRT: 428 AA.
AC 09RGX7:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE MACROPHAGE ACTIVATING LIPOPROTEIN-404 PRECURSOR.
MALP.
Mycoplasma fermentans.
Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
Mycoplasmataceae; Mycoplasma.
NCBI_TaxID=2115;
[1]
SEQUENCE FROM N.A.
RP STRAIN=II-29/1;
RC MEDLINE=99115554; PubMed=9916088;
RX Calcutt M.J., Kim M.F., Karpas A.B., Muhlradt P.F., Wise K.S.;
RA "Differential posttranslational processing confers intraspecies
RT variation of a major surface lipoprotein and a macrophage-activating
RT lipopeptide of Mycoplasma fermentans.";
RL Infect. Immun. 67:760-771(1999).
DR EMBL; AF099209; AAD16393.1; -.
DR InterPro; IPR003760; Bmp.
DR Pfam; PF02608; Bmp; 1.
KW Signal; Lipoprotein.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 428 MACROPHAGE ACTIVATING LIPOPROTEIN-404.
SQ SEQUENCE 428 AA; 47960 MW; D6B69C0BA969CE1 CRC64;

```

```

Query Match 17.1%; Score 393.5; DB 2; Length 428;
Best Local Similarity 26.8%; Pred. No. 6.2e-15;
Matches 127; Conservative 97; Mismatches 178; Indels 71; Gaps 21;

```

```

OY 1 MKRKIKNFKLGLVFPISAIATISAGCWDKETTKEKSADNQNKOITPVSKISGLVNE 60
DB 1 MKRSKR--ILGLSPIALILPAVAVSCG-----NNDSEISKEKDIIYTTN--ANE 50
OY 61 RKSEIMAAKADANKHGLMAIYTAGVYNDNSFNQSSWEALIOQLALNGEITTSVDSST 120
DB 51 KQ---VVKNAELK---LKPVLITDEGKIDKSFNQSAPFALKAIKQVIEINNEPS- 103
OY 121 AELEGRYSILANTNKNWVYLSGFQHGDAFTRMUKIPEKNOLFTEKNIILIGIMTWTENV 180
DB 104 SNFESAINSALSAGHKIWLNGFKHQOSIKQY--IDAHRERLERNQIKIKGIDF-DIE 160
OY 181 IPTGRVINTLYTEAGWLAGYANASFLAKFPSPDTPKRSATVYIGGISPAVDTFAGYLAG 240
DB 161 YKM--FYSLOFINKESEFTTGVAIASWLSQ---DESKRVASFGGAPFGVTTFEGRA 215
OY 241 AGIKAMNLSKDKRTKI-TTDKIEINLGFVDOTSTKERLEQIAS-----KDKPS 289
DB 216 KGIITYVNOKH--KSKKIYHTSPVKLDSG-----TAGCKMNVYINNVLSSTPADVYVNP 268
OY 290 TLLAVAGPTEIFSDI-IANODRVLIGVDTOOLVYTKNKFPTSLKNGSVSVLS 349
DB 269 VILSVAGPAT--FETVRLAANGQYIVGDSOGMT--ODKDRILTSVLKHIOQAVYETLL 324
OY 350 DL-----YTKKSNSRLAGFEFGKKSATVYLGIDKDFVDIADTSLGSDNKL-AT 398
DB 325 DLLEKEEGKYPYVVKDKKADKMSHFQTQ-----KEWIGVAENHNSFTEQOAKIN 376
OY 399 EAISEAKKEFEKTKTIPAEVRKTLPEMPDKOPKOQESLDKLTIDINK 450

```

Db 377 NKIKKAIKMFKE-----LPEDFVKYINSDKALKDGNKIDNVSRLEAIIISAINK 425

RESULT 10
09RGX3 PRELIMINARY; PRT; 428 AA.

AC 09RGX3:
DB 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE MACROPHAGE ACTIVATING LIPOPROTEIN-404 PRECURSOR.
GN MALP.
OS Mycoplasma fermentans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2115;
RN [1]
RP SEQUENCE FROM N.A.
RT STRAIN-K7.
MEDLINE-99115554; PubMed-9916088;
RA Calcutt M.J., Kim M.F., Karpas A.B., Muhlradt P.F., Wise K.S.;
RT "Differential posttranslational processing confers intraspecies
variation of a major surface lipoprotein and a macrophage-activating
lipopeptide of Mycoplasma fermentans.";
RL Infect. Immun. 67:760-771(1999).
DR EMBL; AF099213; AAD16397.1; -.
DR InterPro; IPR003760; Bmp.
DR Pfam; PF02608; Bmp; 1.
KW Signal; Lipoprotein.
FT CHAIN 1 24 POTENTIAL.
FT SIGNAL 1 24 MACROPHAGE ACTIVATING LIPOPROTEIN-404.
SQ SEQUENCE 428 AA; 47875 MW; C7A536B409A60132 CRC64;

Query Match 16.9%; Score 389.5; DB 2; Length 428;
Best Local Similarity 26.8%; Pred. No. 1e-14;
Matches 126; Conservative 99; Mismatches 176; Indels 69; Gaps 20;

Db 5 IKWNFFLGVLFPISAI-ARTISACMDKETTKEKSADNONKQITDVSIGLVNERKS 63
1 MKSKKILGL-SPIAAVLPAAVASCNNDES-----NISFEKDISKYTTTANAGK 52
64 EIMAKADANHFGLMAIVTAGTVNDNSFNQSSWEAIIQOLGALTGEITSVDSSTAE 123
53 VKNAEL-----LKLKPVLTIDESKIDKSFNQSFAELKAIKNOTGIEINSVEPS-SNF 106
124 EGYSSSLANTKKNVYLSGFGHDAFTRWLKIPEKQOLFTEKNIILIGIDMTDTENVIPT 183
107 ESAVNSALSAGHKIIVNLGFKHQOSIKOY--IDAHRLEERNOQIITIGIDF-DIETEKW 163
184 GRYINLYKTEBAGLAGYANASFLAKKFPSPDPTKRSIAIVIGGIGSPAVTDFIAGYLAGI 243
164 --FYSLOFNKESAFETTGATIASWLSQ--DESKRVVASFGVAFPGVTTFNEGFAGGI 218
244 KAMNKNDSKDKTKI-TTQKIEINLGFVDVDTSTKERLEQIAS-----KDKPSTLL 292
219 LYINOKH--KSKKIYHTSPVKLDSG-----TAGKMTVYINNVLSSTPADVKYKPNHYIL 271
293 AVAGPLTEIFSDIIANQDRYLIGVDQSLVYTKKKEFTSILKNLGVSVFVSLDL- 351
272 SVAGPAT--FETVRLANKGQYVIGVSDQGM--QDKRILITSVLKHKAQAVETFLDLI 327
352 -----YTKKSNRNLAFEGFGKKSATYVYLGIDRFVADIADTSLGDNKRL-ATEAI 401
328 LEKEBGYPYVVKDKKAKKWSHFQTO-----KEKWIGVAENHFSNTEQAKINNKI 379
402 SEAKKEFEKTKTIPAEVYKTLLEIPMPD-KOPDKQOESLDKLTITDINK 450
380 KEAIKMFKE-----LPEDFVKYINSDKALKDGNKIDNVSRLEAIIISAINK 425

RESULT 11
09R3N6 PRELIMINARY; PRT; 350 AA.

Db 377 NKIKKAIKMFKE-----LPEDFVKYINSDKALKDGNKIDNVSRLEAIIISAINK 425

RESULT 10
09RGX3 PRELIMINARY; PRT; 428 AA.

AC 09RGX3:
DB 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE MACROPHAGE ACTIVATING LIPOPROTEIN-404 PRECURSOR.
GN MALP.
OS Mycoplasma fermentans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2115;
RN [1]
RP SEQUENCE FROM N.A.
RT STRAIN-K7.
MEDLINE-99115554; PubMed-9916088;
RA Calcutt M.J., Kim M.F., Karpas A.B., Muhlradt P.F., Wise K.S.;
RT "Differential posttranslational processing confers intraspecies
variation of a major surface lipoprotein and a macrophage-activating
lipopeptide of Mycoplasma fermentans.";
RL Infect. Immun. 67:760-771(1999).
DR EMBL; AF099213; AAD16397.1; -.
DR InterPro; IPR003760; Bmp.
DR Pfam; PF02608; Bmp; 1.
KW Signal; Lipoprotein.
FT CHAIN 1 24 POTENTIAL.
FT SIGNAL 1 24 MACROPHAGE ACTIVATING LIPOPROTEIN-404.
SQ SEQUENCE 428 AA; 47875 MW; C7A536B409A60132 CRC64;

Query Match 16.9%; Score 389.5; DB 2; Length 428;
Best Local Similarity 26.8%; Pred. No. 1e-14;
Matches 126; Conservative 99; Mismatches 176; Indels 69; Gaps 20;

Db 5 IKWNFFLGVLFPISAI-ARTISACMDKETTKEKSADNONKQITDVSIGLVNERKS 63
1 MKSKKILGL-SPIAAVLPAAVASCNNDES-----NISFEKDISKYTTTANAGK 52
64 EIMAKADANHFGLMAIVTAGTVNDNSFNQSSWEAIIQOLGALTGEITSVDSSTAE 123
53 VKNAEL-----LKLKPVLTIDESKIDKSFNQSFAELKAIKNOTGIEINSVEPS-SNF 106
124 EGYSSSLANTKKNVYLSGFGHDAFTRWLKIPEKQOLFTEKNIILIGIDMTDTENVIPT 183
107 ESAVNSALSAGHKIIVNLGFKHQOSIKOY--IDAHRLEERNOQIITIGIDF-DIETEKW 163
184 GRYINLYKTEBAGLAGYANASFLAKKFPSPDPTKRSIAIVIGGIGSPAVTDFIAGYLAGI 243
164 --FYSLOFNKESAFETTGATIASWLSQ--DESKRVVASFGVAFPGVTTFNEGFAGGI 218
244 KAMNKNDSKDKTKI-TTQKIEINLGFVDVDTSTKERLEQIAS-----KDKPSTLL 292
219 LYINOKH--KSKKIYHTSPVKLDSG-----TAGKMTVYINNVLSSTPADVKYKPNHYIL 271
293 AVAGPLTEIFSDIIANQDRYLIGVDQSLVYTKKKEFTSILKNLGVSVFVSLDL- 351
272 SVAGPAT--FETVRLANKGQYVIGVSDQGM--QDKRILITSVLKHKAQAVETFLDLI 327
352 -----YTKKSNRNLAFEGFGKKSATYVYLGIDRFVADIADTSLGDNKRL-ATEAI 401
328 LEKEBGYPYVVKDKKAKKWSHFQTO-----KEKWIGVAENHFSNTEQAKINNKI 379
402 SEAKKEFEKTKTIPAEVYKTLLEIPMPD-KOPDKQOESLDKLTITDINK 450
380 KEAIKMFKE-----LPEDFVKYINSDKALKDGNKIDNVSRLEAIIISAINK 425

RESULT 11
09R3N6 PRELIMINARY; PRT; 350 AA.

DT 01-JUN-2001 (Tremblrel. 17, last sequence update)
 DE PUTATIVE SUBSTRATE BINDING LIPOPROTEIN PRECURSOR OF AN ABC
 DE TRANSPORTER.
 GN BMAP.
 OS Clostridium sticklandii.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 NCBI_TaxID=1511;
 OX NCBI_TaxID=1511;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SF370 / ATCC 700294 / SEROTYPE M1;
 RX MEDLINE=21192684; Pubmed=11296296;
 RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
 RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lal H.S., Lin S.P.,
 RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
 RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
 RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
 RT Arch. Microbiol. 175:8-18(2001).
 RL EMBL; AJ276209; CAC14304.1; -
 DR InterPro; IPR003760; Bmp.
 DR Pfam; PF02608; Bmp; 1.
 DR LipoProtein; Complete proteome.
 KW LipoProtein.
 SO SEQUENCE 350 AA; 36396 MW; 7B5A6A95E0BFBD2 CRC64;

Query Match 10.1%; Score 232.5; DB 16; Length 350;
 Best Local Similarity 26.3%; Pred. No. 7.7e-06;
 Matches 113; Conservative 57; Mismatches 153; Indels 107; Gaps 20;

QY 9 KFIIGLVFPLSLATIS-AGCWDKERTKEKSDNONKQITDYSKISGLVNEKKSIMA 67
 DB 4 KFIIGLVFPLSLATIS-AGCWDKERTKEKSDNONKQITDYSKISGLVNEKKSIMA 67
 QY 68 AKADANKHFGALNMAIVTAGTVDNDSFNOSWEAIOQLGALTG-GETTSVDSSTAELEGK 126
 DB 27 SKGASGKRLKAMVMTDGTGVDKSPNOSAMEGLOSGKEMGOKTGTGDFPSTSESE 86
 QY 127 YSSLATNKNVWVLSGFQ-HGDAFTRWLKIPEKOLFTENKIIIGIDMTDENVIPTG 184
 DB 87 YA-TNLDPTAVSGGQILYIGIFA-LKDAIAKAGDNEGVFTIID-DIIRGKD 136
 QY 185 RYINLYTTEAGMLAGYANASFLAKKFPSPDKRSAL-VIGGISPAVDFITAGIAGI 243
 DB 137 NVASVTFPADHEAAYLAGIAAK-TTKTKTVGFGVSGEGTIVITREKGFPAAGV 187
 QY 244 KANMLKNSDKRTKITTDEIKINLGF-DVODSTKERLEQIASKDKPSTLLAVAGPL 298
 DB 188 KS-VDDTIOVKYDAGSPDAAKGTIAAQAAGAD-VIYAAAGTG 233
 QY 299 TELPSDIIANONDR-YLIGVDTQ-SLVTY-KTKNKFSTILKNGSVFS 346
 DB 234 AGVFNEAKAINEKSEADKTVWIGVDRDQKDEKRTSKDGEANFVLASSIKVEKAVOL 293
 QY 347 VLSDLTTRKNSNINAGFERGKKSATVYLGIDRFVDIADTSLEGDKLATEISAKK 406
 DB 294 INKOVADKK-FPGK-TTVY-GIKDGVETIATF-NVSKAEVKAIKEKA 338
 QY 407 EFEKTKTIP 416
 DB 339 KIKSGDIKVP 348
 RESULT 13
 Q9EV91 PRELIMINARY; PRT; 349 AA.
 AC Q9EV91;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, last sequence update)

DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
 DE PUTATIVE SUBSTRATE BINDING LIPOPROTEIN PRECURSOR OF AN ABC
 DE TRANSPORTER.
 GN BMAP.
 OS Clostridium sticklandii.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 NCBI_TaxID=1511;
 OX NCBI_TaxID=1511;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 519T.
 RX MEDLINE=21089007; Pubmed=11271425;
 RA Graentzdorfer A., Pich A., Andreessen J.R.;
 RT "Molecular analysis of the grd-operon encoded proteins of the glycine
 RT reductase and thioresdoxinsystem from Clostridium sticklandii.";
 RL Arch. Microbiol. 175:8-18(2001).
 DR EMBL; AJ276209; CAC14304.1; -
 DR InterPro; IPR003760; Bmp.
 DR Pfam; PF02608; Bmp; 1.
 DR LipoProtein.
 KW LipoProtein.
 SO SEQUENCE 349 AA; 37083 MW; 53AE45114FCD5B2E CRC64;

Query Match 10.0%; Score 229.5; DB 2; Length 349;
 Best Local Similarity 25.6%; Pred. No. 1.1e-05;
 Matches 112; Conservative 66; Mismatches 145; Indels 115; Gaps 22;

QY 1 MKKIKNNKFLGLVPLSLATISAGCWDKERTKEKSDNONKQITDYSKISGLVNE 60
 DB 1 MKKIKNNKFLGLVPLSLATISAGCWDKERTKEKSDNONKQITDYSKISGLVNE 60
 QY 61 KKSIMAKADANKHFGALNMAIVTAGTVDNDSFNOSWEAIOQLGALTGGETTSVDS-S 119
 DB 39 TAAE-----SLKISMVTDVGVKQDSNOSAMELEKAKKDLGIEIGIESKQ 86
 QY 120 TAELEGYSILAN-TKNVWVLSGFQHGDAFTRWLKIPEKOLFTENKIIIGIDMTDTF 177
 DB 87 DADYEPNLETLVDENDLIMGV-GFKMDAI-LAAGN-YPEOKYALIDNDYADAT 139
 QY 178 -ENVIPTGRYINLYTTEAGMLAGYANASFLAKKFPSPDKRSAL-VIGGISPAVDF 235
 DB 140 PDNV-----LGVLFKEEPSYLVG-LIAKM-QSNKIGTIGMDVPIVNR 184
 QY 236 IAGYIAGIKANMLNSDKRTKITTDEIKINLGFVODSTKERLEQIASKDKPSTLLAVA 295
 DB 185 RFGFLAGYAKAN-PDAV-VDOV-FVNAFPDPAKGRAYA 219
 QY 296 GPLTEIFSDII-ANONDRYLIGVDTQSLVYTKNKFSTILNGLG 341
 DB 220 NQMTASGVDTLFFHAAGDSNGAIEPAKEQNKNWIGVDRDQ-DLAPDNVITSAYKRV 276
 QY 342 YSVFSVLSDLTTRKNSNINAGFERGKKSATVYLGIDRFVDIADTSLEGDKLATEAI 401
 DB 277 NAMFNVAKEI-KDGNFAG-GTIVVGLAEGVDIAPTSK-NVPADILEFV 324
 QY 402 SEAKKEFEKTKTIPAE 419
 DB 325 EAEKQIINKKEITVPGTE 342

RESULT 14
 Q97RHO PRELIMINARY; PRT; 350 AA.
 AC Q97RHO;
 DT 01-OCT-2001 (Tremblrel. 18, Created)
 DT 01-OCT-2001 (Tremblrel. 18, last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
 DE LIPOPROTEIN.
 GN SP0845.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;

[1]
RN SEQUENCE FROM N.A.
RP STRAIN-TIGR4;
RC MEDLINE-21357209; PubMed-11463916;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
Peterson S., Heidelberg J., Deboy R.T., Haft D.H., Dodson R.J.,
Durkin A.S., Gwin M., Kolony J.F., Nelson W.C., Peterson J.D.,
Rumayor L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
Holtapple E., Khouri H., Wolf A.M., Uterback T.R., Hansen D.,
McDonald L.A., Feldblyum T.V., Angiuoli S., Dickson T., Hickey E.K.,
Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
R.A. Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.,
"Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae";
RL Science 293:498-506(2001).
DR EMBL: AE007390; AAK74976.1; -
DR TIGR: SP0845; -
DR InterPro: IPR001114; Adenylsucc_synth.
DR InterPro: IPR003760; Bmp.
DR InterPro: IPR00217; Tubulin.
DR Pfam: PF02608; Bmp; 1.
DR PROSITE: PS00227; TUBULIN; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 350 AA; 36746 MW; CBD767B7FF501E69 CRC64;

Query Match 9.7%; Score 223; DB 16; Length 350;
Best Local Similarity 24.8%; Pred. No. 2.7e-05;
Matches 110; Conservative 63; Mismatches 146; Indels 122; Gaps 22;
OY 1 MKKTKMKNKFLGLVFPPLSAITIS--ACQWDEKTTKEKKSADNOKQITDVSKISGLVN 59
DB 1 MKKR-----OMLGLGV-----AVAAVGLAACGRSSRNASSSDVTKT----- 39
OY 60 EKSEIMAKADANKHFGMLNMAIVTAGTVNDNSFNQSMQWALQOLGA---LTGGETSV 116
DB 40 -----AAIVDTGVDKSFQNSMEGLQAMGKHNLSKONGFTY 79
OY 117 DSTAELE-----GKYSILANTNKNVWVLSGFQHDGDAFTRMKIPENKOLFTEKNI 167
DB 80 FQSTEDADVANNLQQAAGSYNLI FG-----GFLNNA-----VDAKKEHTDLNV 125
OY 168 IILGIDMTDTEVNIPTGRITNLTYTEEAGWLAGYANASFLAKKPPSDPTKSAIYIGG 227
DB 136 VALDIVIDKQNV-----ASVTFADNESGYLAGVAAAK-----TTKROYGFVG 170
OY 228 I-SPATDFEYLAGIKAMNLSKSDKTKITTDKIEINLGFVDVDTSKERLEQIASKD 286
DB 171 IESVVISRFEAGFKAGV-----ASVDPSIKVOVDYAG-SFG-DAKKGKTIIAAQYAAAGD 223
OY 287 KPTSLLAAGPL-TEIFSDIIANQDR-----YLIGVDTQSL-VYT---KTKNKEF 333
DB 224 ---IYYQVAGTGAGVFAEAKSLNESPENKRYWVIGVDRDQEAEGKYSKDGKESNFVL 280
OY 334 TSLKNLGVSVFSLDLYTKKSNRNLAGFEPGKSAIYVIGIDRFVDIADTSLSEGD 393
DB 281 VSLKQVGTIVDI-----SNKAERGEFPGQ---VIYVSLDKGVDLAVTNLSEGG 329
OY 394 KRLATEAISEAKKEFEKTKTIP 416
DB 330 KR-----AVEDAKAKILDSGVKVP 348

RESULT 15
O9CFM9 PRELIMINARY; PRT: 350 AA.
AC O9CFM9;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
DE BASIC MEMBRANE PROTEIN A.
GN BMA.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Lactococcus.
OX NCBI_TaxId=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TIGR4;
RA MEDLINE-21235186; PubMed-11337471;
RA Bolotin A., Wincker P., Mauer S., Jallion O., Malarne K.,
Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis asp. lactis I11403";
RL Genome Res. 11:731-753(2001).
DR EMBL: AE006374; AAK0536.1; -
DR InterPro: IPR003760; Bmp.
DR Pfam: PF02608; Bmp; 1.
KW Complete proteome.
SQ SEQUENCE 350 AA; 36652 MW; 2A430D8EEDD4C814 CRC64;

Query Match 9.0%; Score 207; DB 16; Length 350;
Best Local Similarity 24.2%; Pred. No. 0.00022;
Matches 107; Conservative 62; Mismatches 156; Indels 118; Gaps 20;
OY 1 MKKTKMKNKFLGLVFPPLSAITIS--ACQWDEKTTKEKKSADNOKQITDVSKISG 56
DB 1 MKKR-----VIAVSALALASVAVALACCRSHDAAGSGK----- 32
OY 57 LVNERSEIMAKADANKHFGMLNMAIVTAGTVNDNSFNQSMQWALQOLGALTG-----G 111
DB 33 -----AKTD-----LKAIVTEIGVNDNSFNQSMQWALQOLGALGKGTG 75
OY 112 EITSVDSSTAELEGKYSILANTNKNVWVLSGFQHDGDAFTRMKIPENKOLFTEKNIILG 171
DB 76 YTFQSNASASDPTTVNNSAEQGYKLLFGTSLDQTSAAK---NN---PKSNFVIYD 129
OY 172 IMTDTENVIPGRITNLTYTEEAGWLAGYANASFLAKKPPSDPTKSAI-VIGGCI SP 230
DB 130 SVIKQDKNV-----ASATFADNESAYLAGVAAK-----ATKTKKIGFIGQMOSD 174
OY 231 AVTDFEYLAGIKAMNLSKSDKTKITTDKIEINLGFVDVDTSKERLEQIASKDPST 290
DB 175 VITREPKGYEAKKSVN-----PDIKVDVYAGSFSDAKGKTIIAAAMYAGADV 224
OY 291 LLAAGPL-TEIFSDIIA-----NQDR-YLIGVDTQSL-VYT---KTKNKEFTSL 337
DB 225 VYOCAGGVGTGFSEAKALNSTKNEADKRYWVIGVDDQYELGKYSKDGKDSNFVLVSTI 284
OY 338 KNLGVSVFSLDLYTKKSNRNLAG--FEPGKSAIYVIGIDRFVDIADTSLSEGD 395
DB 285 KEVG-----NVVKDIADKTKDGKFPGGTIVYDLKNGVNLGL-----DSANSEIK----- 330
OY 396 LATEAISEAKKEFEKTKTIPAE 418
DB 331 ---DAVAKAKADILDGKITVPSK 350

Search completed: July 18, 2002, 15:31:16
Job time: 123 sec

THIS PAGE BLANK (USPTO)